







Query Match	74.9%	Score 1052.6;	DB 10;	Length 1444;
BASE COUNT	412 a	270 c	342 g	420 t
ORIGIN				
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a></p> <p>Series: IRAC Plate: 11 Row: a Column: 2</p> <p>This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6679586.</p> <p>Location/Qualifiers</p> <p>1..1444</p> <p>/organism="Mus musculus"</p> <p>/db_xref="taxon:10090"</p> <p>/map="FVB/N"</p> <p>/clone="MGC:6226 IMAGE:3592802"</p> <p>/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."</p> <p>/clone_lib="NCI CGAP_Maml"</p> <p>/lab_host="DH10B"</p> <p>/note="vector: pCMV-SPORT6"</p> <p>196..813</p> <p>/codon_start=1</p> <p>/product="RAB1, member RAS oncogene family"</p> <p>/protein_id="AA02077.1"</p> <p>/db_xref="GI:12805233"</p> <p>/db_xref="LOCUSID:5861"</p> <p>/translation="WSSMNEPYDYLFLLLLIGDSGVKSCILLRFADDTYESISTTGVDKQIEIDRTASENVKILVGNKCDLITTKKVDYTTAKERADSLGIFLETSSANAKNVVQESFMMAAEIKRRMGPGATAGCAEKSNVKIGSTPVKSGGGCC"</p>			
REMARK	<p>NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a></p> <p>Contact: MGC help desk</p> <p>Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a></p> <p>Tissue Procurement: Gilbert Smith, Ph.D.</p> <p>cDNA Library Preparation: Life Technologies, Inc.</p> <p>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305</p> <p>Web site: <a href="http://www.shgc.stanford.edu">http://www.shgc.stanford.edu</a></p> <p>Contact: (Dickson, Mark) <a href="mailto:mcgdpaxil@stanford.edu">mcgdpaxil@stanford.edu</a></p> <p>Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.</p>			
COMMENT				
REFERENCE	<p>1 (bases 1 to 1444)</p> <p>Strausberg, R.</p> <p>Direct Submission</p> <p>Submitted (31-Jan-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA</p>			
AUTHORS				
TITLE				
JOURNAL				
ORGANISM	<p>Mus musculus</p> <p>house mouse.</p> <p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.</p>			
ACCESSION	<p>BC002077</p> <p>1444 bp mRNA linear ROD 07-AUG-2002</p> <p>Mus musculus, RAB1, member RAS oncogene family, clone MGC:6226</p>			
VERSION	<p>BC002077</p> <p>GI:12805232</p>			
KEYWORDS	<p>MGC.</p>			
LOCUS	<p>BC002077</p> <p>1444 bp mRNA linear ROD 07-AUG-2002</p> <p>Mus musculus, RAB1, member RAS oncogene family, clone MGC:6226</p>			
DEFINITION	<p>Mus musculus, RAB1, member RAS oncogene family, clone MGC:6226</p>			
DB	<p>1178</p> <p>CTAGCTAACCAACACATGGAAGATATGTTTAAATAATTTGTACTATATGCAATAAAA</p> <p>1319</p> <p>CTAGCTAACCAACACATGGAAGATATGTTTAAATAATTTGTACTATATGCAAGATAAAT</p> <p>1237</p>			
QY	<p>1320</p> <p>AAAA</p> <p>1323</p>			
DB	<p>1238</p> <p>CAAA</p> <p>1241</p>			
RESULT 3	<p>BC002077</p> <p>1444 bp mRNA linear ROD 07-AUG-2002</p> <p>Mus musculus, RAB1, member RAS oncogene family, clone MGC:6226</p>			

	Best Local Similarity	88.3%	Pred. No. 1.9e-160	Matches 1256	Conservative	0	Mismatches	95	Indels	108	Gaps	6
QY	16	GC	GGC	CGC	GCT	GC	AT	TG	TG	TG	TG	75
Db	27	GC	GGC	GGC	GCT	GC	AT	TG	TG	TG	TG	85
QY	76	AC	AGC	CTA	TCA	TCT	CTT	CTT	CGA	TAC	CCG	135
Db	86	AC	AGC	TTA	GC	TAT	CTC	CTC	CTT	CA	TAC	145
QY	136	GC	GGC	CAG	CAG	GGC	-----	GGC	GGT	GGC	GGC	189
Db	146	GC	GT	CAG	CAG	GGC	GGT	GGT	GGC	GGC	GGC	205
QY	190	TG	AAT	CCCGA	TAT	GAT	TAT	TAT	TAT	TAT	TAT	249
Db	206	TG	AAT	CCCGA	TAT	GAT	TAT	TAT	TAT	TAT	TAT	265
QY	250	AG	CTT	GC	CTT	CTT	CTT	AG	TTG	CAG	TAT	309
Db	266	AG	CTC	TG	CTT	CTC	TAT	AG	TTT	GC	AAT	325
QY	310	TT	GC	T	GT	GT	GT	GT	GT	GT	GT	369
Db	326	TT	GC	T	GT	GT	GT	GT	GT	GT	GT	385
QY	370	TA	-----	-----	-----	-----	-----	-----	-----	-----	-----	371
Db	386	TA	T	AT	GG	AC	C	A	G	C	A	445
QY	372	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	393
Db	446	CC	CA	T	GC	A	T	C	A	T	A	505
QY	394	AG	T	GC	T	GC	A	T	T	GC	A	453
Db	506	AG	T	GC	T	GC	A	T	T	GC	A	565
QY	454	ACA	A	T	GT	GT	GT	GT	GT	GT	GT	513
Db	566	ACA	A	T	GT	GT	GT	GT	GT	GT	GT	625
QY	514	AT	T	CC	T	T	GC	A	T	T	GC	573
Db	626	AT	T	CC	T	GC	A	T	T	GC	A	685
QY	574	CT	T	C	A	T	G	A	C	A	T	633
Db	686	CT	T	C	A	T	G	A	C	A	T	745
QY	634	GT	CT	G	A	AG	T	T	A	A	T	693
Db	746	GT	CT	G	A	AG	T	T	A	A	T	805
QY	694	GC	T	C	T	A	A	A	T	T	G	753
Db	806	GC	T	C	T	A	A	A	T	T	G	865
QY	754	TG	A	A	A	A	A	A	A	A	A	813
Db	866	TG	A	A	A	A	A	A	A	A	A	925
QY	814	GT	C	T	C	C	A	A	A	A	A	872
Db	926	GT	T	T	T	T	T	T	T	T	T	984
QY	873	AC	T	C	A	A	G	C	T	A	A	932
Db	985	AC	T	C	A	A	G	C	T	A	A	1044
QY	933	A	A	T	A	T	G	T	A	A	T	992



Db	1045	AAATATGTGTATACCTCTGCTTCTCTGATACCAAGATGTTCCCGTGGTTAG	1104	/gene="DKFZp564B163"
QY	993	AAATATATTTTGTGATGTTATATATGCAATGTTAGATGTTAGTCTTGAA	1052	/gene="DKFZp564B163"
Db	1105	AAATATATTTTGTGATGTTATATATGCAATGTTAGATGTTAGTCTTGAA	1164	/note="Rabl, splice variant"
QY	1053	GATGAATTCAGCAATTTTGTATCAACAGCAAGAGTGTCTGACTTTCATGCAT	1112	/codon_start=1
Db	1165	GATGAATTCAGCAATTTTGTATCAACAGCAAGAGTGTCTGACTTTCATGCAT	1224	/product="hypothetical protein"
QY	1113	AAATTTAGTGAATGTTATATGATAGATCTGATTTGCTGCTTCTGATAGATTAT	1172	/protein_id="CA43369.1"
Db	1225	AAATTTAGTGAATGTTATATGATAGATCTGATTTGCTGCTTCTGATAGATTAT	1284	/db_xref="GI:4886443"
QY	1173	AAATGAAGATTAACATCTGATTAATAGTTCCTGACACCTGCAATATATTTGG	1232	/translation="MSSMPEDYLFKLLIDSGVSKCLLPADPTYESYIST
Db	1285	AAATGAAGATTAACATCTGATTAATAGTTCCTGACACCTGCAATATATTTGG	1344	GVDKIRITLIDGKTIKLIEFADSLGIPLETSKATNATVDSFMTAAEIKRMGP
QY	1233	CTGAGAAATTTGATTTGTTGACACTATGTAACAAACATGATATGTTTAT	1292	GATGCAKSNVRIQSTPVGKGGCC"
Db	1345	CTGAGAAATTTGATTTGTTGACACTATGTAACAAACATGATATGTTTAT	1401	polya_signal
QY	1293	AAATATTTGCTTATTTGAACTAAAAAATTTTAAAAA	1335	polya_site
Db	1402	AAATATTTGCTTATTTGAACTAAAAAATTTTAAAAA	1444	BASE COUNT 344 a 213 c 279 g 372 t
RESULT 4				ORIGIN
LOCUS	HSB800023	1208 bp	mRNA	linear
DEFINITION	Human sapiens mRNA: CDNA DKFZp564B163 (from clone DKFZp564B163);			
ACCESSION	AL050268			
VERSION	AL050268.1	GI:4886442		
KEYWORDS	human.			
SOURCE	human.			
ORGANISM	Human sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
JOURNAL	Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glasl S.,			
COMMENT	Ansorge W., Boecker M., Bloeker A., Bauersachs S., Blum H.,			
	Lauder J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,			
	Mewes H.W., Ottenwaelder B., Oberaler B., Tampe J., Heubner D.,			
	Wambutt R., Korn B., Klein M., and Poustka A.			
	Toward a Catalog of Human Genes and Proteins: Sequencing and			
	Analysis of 500 Novel Complete Protein Coding Human CDNA			
	Genome Res. 11 (3), 422-435 (2001)			
JOURNAL	11230166			
PUBMED	2 (bases 1 to 1208)			
REFERENCE	Wambutt R., Heubner D., Mewes H.W., Gassenhuber J. and Wiemann S.			
AUTHORS	Direct Submission			
JOURNAL	Submitted (10-MAR-1999) WIPs, Am Klopfersplitz 18a, D-82152			
COMMENT	Martinsried, GERMANY			
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer			
	Research Center (DKFZ). Email s.wiemann@dkfz-heidelberg.de;			
	consortium of the German Genome Project.			
	This clone (DKFZp564B163) is available at the RZPD in Berlin.			
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059			
	Berlin-Charlottenburg, GERMANY. Email: clonerzpd.de			
	Information about the clone and the sequencing project is available			
	at http://www.mips.biochem.mpg.de/proj/cdna/.			
FEATURES	Location/Qualifiers			
SOURCE	1..1208			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="DKFZp564B163"			
	/tissue_type="brain"			
	/clone_id="564 (synonym: hfbz2). Vector pAMP1; host			
	X1-2blue; sites NotI + SalI			
	/dev_stage="fetal"			
gene	191..580			

DB	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
OY	780	ACTGATATCTACCTGCATACCAACAGATCTTACCGCTCCGACCAAAAGGTCAGAGATTGTA									
Db	659	ACTGATATCTACCTGCATACCAACAGATCTTACCGCTCCGACCAAAAGGTCAGAGATTGTA									
OY	840	ATGCTCACTACTGACTTTTTTTTTTATTCCTTGACTCAAGACAGCTACCTCATTTTCAG									
Db	719	ATGCTCACTACTGACTTTTTTTTTTATTCCTTGACTCAAGACAGCTACCTCATTTTCAG									
OY	900	AACGTTTAAACCTTGTGTGCGGCTTATTAATAATATGTGTAACTCTGTGCTT									
Db	779	AACGTTTAAACCTTGTGTGCGGCTTATTAATAATATGTGTAACTCTGTGCTT									
OY	960	CCTGATACACAGACTGTTCCCGGTGGTGGTGAATATATTTTGTGTGATGTTATAT									
Db	839	CCTGATACACAGACTGTTCCCGGTGGTGGTGAATATATTTTGTGTGATGTTATAT									
OY	1020	GGCATTTTATATGTCACAGCTTACCTCTCTAATAATGAAGTTCAGCCATTGTATCAA									
Db	899	GGCATTTTATATGTCACAGCTTACCTCTCTAATAATGAAGTTCAGCCATTGTATCAA									
OY	1080	CAGCAACAAGCAGTGTCTGTCTACTCTTCATGCAATAAGTTAGTGAAGATTATATGA									
Db	959	CAGCAACAAGCAGTGTCTGTCTACTCTTCATGCAATAAGTTAGTGAAGATTATATGA									
OY	1140	ATCTGATTTCCTAGTCTTCTCTTGTAGAGTTATAATGAAAGATTACATCTGATTA									
Db	1019	ATCTGATTTCCTAGTCTTCTCTTGTAGAGTTATAATGAAAGATTACATCTGATTA									
OY	1200	ATATTTCTCTATCTCGCATATATATTTGGTGGTCAGAAATATGTATTTGTGGACA									
Db	1079	ATATTTCTCTATCTCGCATATATATTTGGTGGTCAGAAATATGTATTTGTGGACA									
OY	1260	CTATGTACCAACAACACTGAAGATATGTTATAATAATTTGACTATTGGAACTAAAA									
Db	1139	CTATGTACCAACAACACTGAAGATATGTTATAATAATTTGACTATTGGAACTAAAA									
OY	1320	AAAAAAAAAA 1329									
Db	1199	AAAAAAAAAA 1208									
RESULT 5											
LOCUS	MYPT1	1428 bp	linear								
DEFINITION	Mouse mRNA for ras-related Yp1 protein.										
ACCESSION	Y00094										
VERSION	Y00094.1										
KEYWORDS	ras-related protein; YP1 gene.										
SOURCE	Mus musculus.										
ORGANISM	Mus musculus										
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.										
AUTHORS	1 (bases 1 to 1428)										
TITLE	Galwitz, D.										
JOURNAL	Submitted (13-NOV-1987) Galwitz D., Max-Planck-Institut fuer biophysikalische Chemie, Abt. Molekulare Genetik, Postfach 2841, D-3400 Goettingen, FRG										
REFERENCE	2 (bases 1 to 1428)										
AUTHORS	Habbruck, H., Disela, C., Wagner, P. and Galwitz, D.										
TITLE	The ras-related Yp1 protein is an ubiquitous eukaryotic protein: isolation and sequence analysis of mouse cDNA clones highly homologous to the yeast YP1 gene										
JOURNAL	EMBO J. 6 (1987) In press										
FEATURES	EMBO J. 6 (1987) In press										
SOURCE	Location/Qualifiers										
	1..1428										
	/organism="Mus musculus"										
	/db_xref="taxon:10090"										
	/cell_line="F9 and C3H10T1/2"										
	130..747										
	/note="Yp1 protein (AA 1-205)"										

```

misc_feature
    1401..1406
        /note="pot. polyA signal"
polyA_site
    1428
        /note="polyA site"
BASE COUNT      410 a      263 c      324 g      431 t
ORIGIN
Query Match      72.1%;   Score 1013.2;  DB 10;   Length 1428;
Best Local Similarity 88.0%;   Pred No. 4,4e-156;
Matches 1204;   Conservative 0;   Mismatches 38;   Indels 106;   Gaps 5;

/db_xref="GI:55457"
/db_xref="MGD:MG1:97842"
/db_xref="SWISS-PROT:P11476"
/translation="WSSNRPDYDLFKLLILGDSGVKSCILLRFPADDTYESYISTT
GVDRKIRTELDGRTIKLQIWDITAGEEFRRTISSYKNGAGIILVYDVTDQSEFNN
KQWOLEIDRYASENNKLLVGNKCDLTKKVDVTTAAEFDSIGIPLFETSAKNATV
VEOSFMVMAAEIKRMRMGATAGAGAEKSNVKGISNPVKGSGGGC"

```

Oy		797	TACAACAAATTCCTTACCAGCTGCTCCACAAGAAGTGCAAGATTGTGAATATGGCAATACTGCAC-T	855
Dd		843	TACAACAAATTCCTTACCAGCTTTCCAC -AAGTCAGAGATTGGTAATGTGCATACTGACTT	901
Oy		856	TTTTTTTTTATTCOCCTTGACTCACAACAGCTAACCTTCATTTCCAGAAAGCTTTTAAACCTT	915
Dd		902	TTTTTTTTTATTCOCCTTGACTCACAACCGCTAACCTTCATTTCCAGAAAGCTTTTAAACCTT	961
Oy		916	TGTCGTGCTGCTGTTTAAAAATATATGTGTAACTCTGTGCTTCCGTATACAGACACTGT	975
Dd		962	TGTCGTGCTGCTGTTTAAAAATATGTGTAACTCTGTGCTTCCGTATACAGACACTGT	1021
Oy		976	TTCOCGTGCTGCTGTTTAAAAATATATGTGTAACTCTGTGCTTCCGTATACAGACACTGT	1035
Dd		1022	TTCOCGTGCTGCTGTTTAAAAATATATGTGTAACTCTGTGCTTCCGTATACAGACACTGT	1081
Oy		1036	AGGTTTAACTCTGTCGAAGTGAAGTGTACGCCATTTGTATCAACACGACAGCATGTC	1095
Dd		1082	GGGTTTAGCTCTTGCAAGATGAAGTGTACGCCATTTGTATCAACACGACAGCATGTC	1141
Oy		1096	TGTCACTCTTCACATGATTAAGTTAGTGAAGATGTATATGTAAAGTCTGATTTGCTAGTT	1155
Dd		1142	TGTCACTCTTCACACCATTAAGTTAGTGAAGATGTATATGTAAAGTCTGATTTGCTAGTT	1201
Oy		1156	CCTTCCTGTAGAGTTATTAATGGAAGATTAACATCTATCTGATTAATAGTTTCCTCACT	1215
Dd		1202	CCTTCCTGTAGAGTTATTAATGGAAGATTAACATCTATCTGATTAATAGTTTCCTCACT	1261
Oy		1216	CTGCATATTAATTTTGTGGCGCAGAAATTTGTATTTGTTCGCACACTATGTAAACAAAACA	1275
Dd		1262	CTGCATATTAATTTTGTGGCGCAGAAATTTGTATTTGTTCGCACACTATGTAAACAAA--NA	1318
Oy		1276	CTGAGAGATATGTTTAAATTAATTTGTACTATTGTGAGACTAAAAAAA	1323
Dd		1319	CTGAGAGATATGTTTAAATTAATTTGTACTATTGTGAGACTAATATATCAA	1366
RESULT 6				
AF226873				
LOCUS	AF226873	2686 bp	mRNA linear	ROD 13-FEB-2000
DEFINITION	Mus musculus small GTP-binding protein RAB1A mRNA, complete cds.			
ACCESSION	AF226873			
VERSION	AF226873.1	GI:96969621		
KEYWORDS				
SOURCE				
ORGANISM	Mus musculus.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
AUTHORS	1. (bases 1 to 2686)			
JOURNAL	Mu.G. and Dorn.G.W. II.			
REFERENCE	Mouse RAB1A, member of RAS oncogene family mRNA Unpublished			
AUTHORS	2. (bases 1 to 2686)			
TITLE	Mu.G. and Dorn.G.W. II.			
JOURNAL	Direct Submission			
FEATURES	Submitted (19-JUN-2000) Internal Medicine, University of Cincinnati, 231 Bethesda Avenue, Cincinnati, OH 45267-0590, USA			
source	Location/Qualifiers			
	1..2686			
	/organism="Mus musculus"			
	/db_xref="taxon:10090"			
	29..646			
	/note="member of RAS oncogene family"			
	/codon_start=1			
	/product="small GTP-binding protein RAB1A"			
	/protein_id="AAP384.1"			
	/db_xref="GI:6969622"			
	/translation="SSSNPEEDYDFKLLIGDSNGSKSLLEFRADDTYESITSTI GVMDIKTRIDSENNKLIVGNRCDLTKRYVYTAAEFSDISGLTFLETSAKNANN VEOSFTNMAELIRKKMGAGTAGAKESNVASTIPVKSGGGCC"			
BASE COUNT	778 a	520 c	563 g	825 t

[illegible]

QY	1075	TCAAACGACCAACAGACGTCTGCACCTTCCATGCAATAAAGTTAGGACATGTTAT	1134
Db	1020	TCACACGACCAACACCATGCTCTGCAGTTCCAGCATTAAGTTAGGACGTTAT	1079
QY	1135	GTAAGATTCGATTTGCTAGTCTTCCCTGTGAGATTAATAATGGAAGATTACACTATCT	1194
Db	1080	GTAAGATTCGATTTGCTAGTCTTCCCTGTGAGATTAATAATGGAAGATTACACTATCT	1139
QY	1195	GATTAATAGTTCTTCATCTCTGCATATATTTGTGGCTGCAGAAATTTGTAATTTGTT	1254
Db	1140	GATTAATAGTTCTTCATCTCTGCATATATTTGTGGCTGCAGAAATTTGTAATTTGTT	1199
QY	1255	GCACACATGATTAACAAAACAACTGAGATATGTTTAATTAATATTGACTATTGGAAGT	1314
Db	1200	GCACACATGATTAACAA--AACTGAGATATGTTTAATTAATATTGACTATTGGAAGT	1256
QY	1315	AAAAAAAA 1323	
Db	1257	AATATCAAA 1265	

RESULT 7	AC123143	LOCUS	AC123143	DEFINITION	AC123143	ACCESSION	AC123143	VERSION	AC123143.2	KEYWORDS	HTGS PHASE1.	SOURCE	Norway rat.	ORGANISM	Rattus norvegicus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	REFERENCE	1 (bases 1 to 151988)	AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C., Aisburooks,S.L., Amaraltinge,H.C., Are,J.R., Ayale,M., Banks,T., Barbalea,J., Benton,J., Blmage,K., Blankenburg,K., Bonnal,D., Bouck,J., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buchak,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dethorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Deederich,D.A., Delaney,R.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabali,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,R., Hart,M., Havlik,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsli,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulsgeid,H., Lozadzo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Mahestrat,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,A., Morris,S., Moser,M., Neel,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunolu,G., Oragunye,N., Oyedelo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojupokan,I., Rolfe,M., Ruliz,S., Saverly,G., Schier,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmami,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.
----------	----------	-------	----------	------------	----------	-----------	----------	---------	------------	----------	--------------	--------	-------------	----------	-------------------	---	-----------	-----------------------	---------	---

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
-----
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Unpublished
Direct Submission
2 (bases 1 to 151988)
Worley, K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 151988)
Worley, K.C.
Direct Submission
Submitted (17-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 4, 2002 this sequence version replaced gi:21239895.
-----
Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: GMGF
Center clone name: CH230-247B24
Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 108541 bases at least Q40
Consensus quality: 113010 bases at least Q30
Consensus quality: 115635 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This is a "working draft" sequence. It currently
* consists of 50 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
1041: contig of 1041 bp in length
*
1042 1141: gap of unknown length
*
1142 2641: contig of 1500 bp in length
*
2642 2741: gap of unknown length
*
2742 3831: contig of 1090 bp in length
*
3832 3931: gap of unknown length
*
3932 3516: contig of 1685 bp in length
*
3517 5716: gap of unknown length
*
5717 6732: contig of 1016 bp in length
*
6733 6833: gap of unknown length
*
6833 8189: contig of 1357 bp in length
*
8189 8290: gap of unknown length
*
8290 9410: contig of 1121 bp in length
*
9411 9510: gap of unknown length
*
9511 10654: contig of 1144 bp in length
*
10655 11847: gap of unknown length
*
11848 11947: contig of 1093 bp in length
*
11948 11948: gap of unknown length
*
11948 13492: contig of 1545 bp in length
*
13493 13592: gap of unknown length
*
13593 14687: contig of 1095 bp in length
*
14688 14787: gap of unknown length
*
14788 16151: contig of 1364 bp in length
*
16152 16251: gap of unknown length
*
16252 18133: contig of 1882 bp in length
*
18134 18233: gap of unknown length
*
18234 19865: contig of 1632 bp in length
*
19866 19965: gap of unknown length
*
19966 21309: contig of 1244 bp in length

```

[illegible]

```

QY      863 TATTCCTGACACAGACAGCTACCTTCATTTTTCAGAACTGTTTAAACCTTGTGTGC 922
Db      76881 TATTCCTGACACAGACAGCTACCTTCATTTTTCAGAACTGTTTAAACCTTGTGTGC 76940
QY      923 TGGTTTAAATTAATGCTGCTATCCTGTTGCTTCCGATATCCAGACTGTTTCCCGT 982
Db      76841 TGGTTTAAATTAATGCTGCTATCCTGTTGCTTCCGATATCCAGACTGTTTCCCGT 77000
QY      983 GGTGGTTAGCAATATATTTGTTTGTGATTTATATTTGCAATGTTT----- 1027
Db      77001 GGTGGTTAGCAATATATTTGTTTGTGATTTATATTTGCAATGTTTGGGTTTGA 77060
QY      1028 ----- 1027
Db      77061 GTTGGGATTTAGCTGACGTGTAGAGCGCTTGCCTACCAAGCGCCCTGGGTTGCG 77120
QY      1028 -----TAGATGTCAGGTTTACGCTTCTGAAGTGAAGTTCAAGCC 1066
Db      77121 TCCCGACCTCGAAAAAATAATAGATGTTGGGTTTACGCTTCTGAAGTGAAGTTCAAGCC 77180
QY      1067 ATTTTGTATCAACACAGCAGACAGCTGCTGCTCACTTCCATGCATTAAGTTTACTGAGA 1126
Db      77181 ATTTTGTATCAACATAGCAGACAGCTGCTGCTCACTTCCATGCATTAAGTTTACTGAGA 77240
QY      1127 TGTATATGTAGATGCTGATTTGCTAGTCTTCTCTGATAGATTAAATGGAAGATTA 1186
Db      77241 CATATATGTAGATGCTGATTTGCTAGTCTTCTCTGATAGATTAAATGGAAGATTA 77299
QY      1187 CACTATGATGATTAATGCTTCTGATCTGCTGATATATTTGGCTGCGAGATTAATGCT 1246
Db      77300 CACTATGATGATTAATGCTTCTGATCTGCTGATATATTTGGCTGCGAGATTAATGCT 77359
QY      1247 AATTTGTGACACACTATGTAACAAACAGACTGAGATATGTTTAAATATTTACTACTA 1306
Db      77360 AATTTGTGACACACTATGTAACAA--AACTGAAGACATGTTTAAATATTTACTACTA 77416
QY      1307 TTGGAAGTAAAAAAA 1323
Db      77417 TCGGAAGTAAATATCAAA 77433

RESULT 8
AC007318/ 204230 bp DNA linear PRI 07-NOV-2001
LOCUS Homo sapiens BAC clone RP11-420C9 from 2, complete sequence.
DEFINITION AC007318
ACCESSION AC007318.4 GI:13435282
VERSION HTG.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 204230)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 204230)
AUTHORS Mollan, A., Hawkins, M., Kallicki, J., Harfins, R. and Lehnert, L.
TITLE The sequence of Homo sapiens BAC clone RP11-420C9
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 204230)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (16-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 204230)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,

```

```

REFERENCE MO 63108, USA
AUTHORS 5 (bases 1 to 204230)
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 6 (bases 1 to 204230)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Mar 23, 2001 this sequence version replaced gi:17630984.
----- Genome Center.
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0420C09
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osseogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Teleno, M., Catalano, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION: The clone sequenced to the left is RP11-350H15; the clone sequenced to the right is AC034121. Actual start of this clone is at base position 1 of RP11-420C9; actual end is at base position 204230 of RP11-420C9.

The sequence RP11-420C9 from base position 44230 to 46506 consists of PCR product from clone DNA. Base position 44700 to 44774 represents sequence of a single plasmid subclone derived from PCR amplification of clone DNA.

#### FEATURES

```

source      1..204230
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="2"
             /map="2"
             /clone="RP11-420C9"
             /clone_id="RPCI-11"
             /rpt_family="Alu"
             /rpt_family="Alu"
             /rpt_family="MER1-type"
repeat_region 179..486
repeat_region 890..1082
repeat_region 1458..1789

```

```
repeat_region 2183..2483 /rpt_family-"Alu"
repeat_region 2573..2874 /rpt_family-"Alu"
repeat_region 3336..3642 /rpt_family-"Alu"
misc_feature 3621..4353 /rpt_family-"Alu"
misc_feature 3656..3996 /note="similar to EST AV725228 (NID:g10830405)"
misc_feature 3687..4148 /note="similar to EST F11955 (NID:g706276)"
misc_feature 3841..4110 /note="similar to EST T65650 (NID:g674695) yc76g03.r1"
misc_feature 4167..4232 /note="similar to EST AV725399 (NID:g10830748)"
misc_feature 4178..4353 /note="similar to EST AV725399 (NID:g10830748)"
misc_feature 4178..4353 /note="similar to EST AA130204 (NID:g1691412) z129b02.r1"
misc_feature 4178..4353 /note="similar to EST AW131666 (NID:g6133273) xf32h07.r1"
misc_feature 4178..4353 /note="similar to EST AW960488 (NID:g6150172)"
misc_feature 4178..4353 /note="similar to EST BF525934 (NID:g11613295)"
misc_feature 4180..4353 /note="similar to EST AT808181 (NID:g5394747) wf93c04.r1"
misc_feature 4654..5066 /note="similar to EST AT278206 (NID:g3900474) q163d03.r1"
misc_feature 4994..5723 /note="similar to EST BF575111 (NID:g11648823)"
misc_feature 5136..5483 /note="similar to EST AT221036 (NID:g3803239) q909a10.r1"
misc_feature 5283..5685 /note="similar to EST BF525934 (NID:g11613295)"
misc_feature 5283..5600 /note="similar to EST AT131666 (NID:g6133273) xf32h07.r1"
misc_feature 5283..5599 /note="similar to EST AT808181 (NID:g5394747) wf93c04.r1"
misc_feature 5283..5437 /note="similar to EST AW960488 (NID:g6150172)"
misc_feature 5283..5335 /note="similar to EST AA130204 (NID:g1691412) z129b02.r1"
misc_feature 5444..5599 /note="similar to EST AV725228 (NID:g10830405)"
misc_feature 5464..5935 /note="similar to EST AT079760 (NID:g3416011) oy43b05.s1"
misc_feature 5467..5919 /note="similar to EST AT802095 (NID:g5367567) tx24g01.r1"
misc_feature 5523..6088 /note="similar to EST AT805795 (NID:g5392361) lj95b10.r1"
misc_feature 5596..6168 /note="similar to EST AA443176 (NID:g2155851) zx98g09.r1"
misc_feature 5608..6176 /note="similar to EST AT637917 (NID:g4690151) tl11h07.r1"
misc_feature 5621..6180 /note="similar to EST AT949154 (NID:g5741464) wq08d02.r1"
misc_feature 5807..6166 /note="similar to EST AW205023 (NID:g6504495)"
misc_feature 5816..6254 /note="similar to EST AW090324 (NID:g6047668) xc66a12.r1"
misc_feature 5940..6169 /note="similar to EST AA478834 (NID:g2207466) zv09g06.r1"
misc_feature 6027..6268 /note="similar to EST D20296 (NID:g501393)"
misc_feature 6038..6470 /note="similar to EST AM391595 (NID:g6896254)"
misc_feature 6053..6275 /note="similar to EST AT803303 (NID:g5366775) tc17d12.r1"
misc_feature 6132..6947 /note="similar to EST AM380429 (NID:g6885088)"
misc_feature 6132..6947 /note="similar to EST AU125475 (NID:g10950191)"
```

```
misc_feature 6663..6949 /note="similar to EST T07351 (NID:g318500)"
misc_feature 6861..7737 /note="similar to EST AU124391 (NID:g10949107)"
misc_feature 7006..7454 /note="similar to EST AA825455 (NID:g2898766) oe64f08.s1"
misc_feature 7130..7930 /note="similar to EST BF242402 (NID:g11156329)"
repeat_region 7171..7467 /rpt_family-"Alu"
misc_feature 7320..7596 /note="similar to EST AV703078 (NID:g10719408)"
misc_feature 7501..8168 /note="similar to EST BE279740 (NID:g9154735)"
misc_feature 7650..7920 /note="similar to EST AV703078 (NID:g10719408)"
misc_feature 7790..8251 /note="similar to EST AA255857 (NID:g1891513) za27h08.r1"
misc_feature 7945..8533 /note="similar to EST AV601747 (NID:g9724069)"
misc_feature 8047..8633 /note="similar to EST AW550283 (NID:g7195711)"
misc_feature 8065..8647 /note="similar to EST AU149130 (NID:g11010651)"

Query Match 58.28; Score 817.2; DB 9; Length 2042230;
Best Local Similarity 98.64; Pred No. 2,1e-122;
Matches 819; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

502 AGAATTGCGATTCCTGCGATTCGTTTGGAAACCGTCGATGAGATGCAAGA 561
10336 AGAATTGCGATTCCTGCGATTCGTTTGGAAACCGTCGATGAGATGCAAGA 10277
562 ATGTGAACACGCTCTTTTANGACGATGCGACGTGATTTAAAAGCGAATGGCTCCGAG 621
10276 ATGTGAACACGCTCTTTTANGACGATGCGACGTGATTTAAAAGCGAATGGCTCCGAG 10217
622 CAACACGCTGCTGCTGAGAGAGTCCAAATGTTAAATTCAGACGCTCCAGCAACGT 681
10216 CAACACGCTGCTGCTGAGAGAGTCCAAATGTTAAATTCAGACGCTCCAGCAACGT 10157
682 CAGGTGAGAGTTCGCTAAATTTGGCTCCATCCTTTTCACAGCAATGAATTTGCA 741
10156 CAGGTGAGAGTTCGCTAAATTTGGCTCCATCCTTTTCACAGCAATGAATTTGCA 10097
742 TCTGAACCAAGTAAAAAACAATTCCTGAAATTCATCTATGATGACGCTACAA 801
10096 TCTGAACCAAGTAAAAAACAATTCCTGAAATTCATCTATGATGACGCTACAA 10037
802 CAGATTCCTAACGCTCCACAAAGTCAAGATTTGAATGGTCAATGACTTTT 861
10036 CAGATTCCTAACGCTCCACAAAGTCAAGATTTGAATGGTCAATGACTTTT 9977
862 TTTATCCCTGACTCAGACGCTAATCTTTCAGAACTGTTAAACCTTTGTG 921
9976 TTTATCCCTGACTCAGACGCTAATCTTTCAGAACTGTTAAACCTTTGTG 9917
922 CTGCTTTTAAATGAATGTGTATATCTTGTGCTTCGATACAGACGCTTTCCG 981
9916 CTGCTTTTAAATGAATGTGTATATCTTGTGCTTCGATACAGACGCTTTCCG 9857
982 TGTGTTGTTAAATGAATGTGTATATCTTGTGCTTCGATACAGACGCTTTCCG 1041
9856 TGTGTTGTTAAATGAATGTGTATATCTTGTGCTTCGATACAGACGCTTTCCG 9797
1042 AGCTTTGGAAGATGAAGTTCAAGCATTTTGTATCAACAGCAACAGCATGCTGCAC 1101
9796 AGCTTTGGAAGATGAAGTTCAAGCATTTTGTATCAACAGCAACAGCATGCTGCAC 9737
1102 TTTCCATGATAAAGTTTATGATAGATTATATGTAAGTCTGATTTGCTAGTTCTCT 1161
9736 TTTCCATGATAAAGTTTATGATAGATTATATGTAAGTCTGATTTGCTAGTTCTCT 9677
```



QY 1162 TGTAGCTTAAATGGAAGATACATATCATTAATAGTTCCTTCATCTGCAT 1221  
 DB 9676 TGTAGCTTAAATGGAAGATACATATCTGTAAATGTTCTTCATCTGCAT 9617  
 QY 1222 ATATTTGTGCTGCAGAAATTTTGTATTTTGTGCACACTATGTAACAACTGAG 1281  
 DB 9616 ATAAATTTGGCTCAGAAATTTGTATTTTGTGCACACTATGTAACAACTGAG 9557  
 QY 1282 ATATTTTAAATTAATTTGACTTATTTGGAAGTAAAAAAA 1323  
 DB 9556 ATATGTTTAAATTAATTTGACTTATTTGGAAGTAAAAAAA 9515

RESULT 9  
 AL606522 203071 bp DNA linear ROD 22-MAY-2002  
 LOCUS Mouse DNA sequence from clone RP23-19L22 on chromosome 11, complete  
 DEFINITION sequence.  
 ACCESSION AL606522 GI:20792609  
 VERSION AL606522.6 GI:20792609  
 KEYWORDS house mouse.  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 203071)  
 AUTHORS Almeida, J.  
 JOURNAL Direct Submission  
 Submitted (22-MAY-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humbrey@sanger.ac.uk  
 On May 14, 2002 this sequence version replaced gi:1968263.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP23-19L22 is from the RP23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see http://www.chori.org/dacpac/home.htm  
 VECTOR: pBAC3.6.

FEATURES  
 source location/Qualifiers  
 1..203071  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone="RP23-19L22"  
 /clone\_1lb="RP23-19L22"  
 /clone\_1lb="RP23-19L22"  
 BASE COUNT 57371 a 44818 c 43421 g 57461 t

ORIGIN  
 Query Match 52.8%; Score 741.4; DB 10; Length 203071;  
 Best Local Similarity 96.2%; Pred. No. 3.5e-110;  
 Matches 792; Conservative 0; Mismatches 26; Indels 5; Gaps 3;

QY 502 AGGAATTTGCTGATTCCTTGGATTCGTTTGGAAACAGTGTCTAGAGTCAACA 561  
 DB 177379 AGGAATTTGCTGATTCCTTGGATTCGTTTGGAAACAGTGTCTAGAGTCAACA 177438  
 QY 562 ATGTAGAACAGCTTTTATGACGATGCGACGTGAGATTAACGGAATGGCTCCGAG 621

DB 177439 ATGTAGAACAGCTTTTATGACGATGCGACGTGAGATTAACGGAATGGCTCCGAG 177498  
 QY 622 CAACAGCTGTGTGCTAGAGATTCATTTTAAATTCACAGCACTCAATCAAGCAGT 681  
 DB 177499 CTACAGCTGTGTGCTAGAGATTCATTTTAAATTCACAGCACTCAATCAAGCAGT 177558  
 QY 682 CAGGTGAGAGTGTGCTAGAGATTCATTTTAAATTCACAGCACTCAATCAAGTGA 741  
 DB 177559 CAGGTGAGAGTGTGCTAGAGATTCATTTTAAATTCACAGCACTCAATTCGAA 177618  
 QY 742 TGTGAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 801  
 DB 177619 TGTGAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 177678  
 QY 802 CAGATTTCTACCGTCTCCACAAAGTGCAGATTTGAATGTCATCTGAC--TTTTT 860  
 DB 177679 CAGATTTCTACCGTCTCCAC--AAGTGCAGATTTGAATGTCATCTGACCTTTT 177737  
 QY 861 TTTATTCCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 920  
 DB 177738 TTTATTCCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 177797  
 QY 921 GCTGCTTTAAATTAATTTGCTGTATCTGTTGCTTCTGTTGCTGTTGCTGTTGCTGTTGCT 980  
 DB 177798 GCTGCTTTAAATTAATTTGCTGTATCTGTTGCTTCTGTTGCTGTTGCTGTTGCTGTTGCT 177857  
 QY 981 GTGGTGTGTTGAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1040  
 DB 177858 GTGGTGTGTTGAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 177917  
 QY 1041 TAGTCTTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1100  
 DB 177918 TAGTCTTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 177977  
 QY 1101 CTTTCATGCTAAAGTTTATGAGATTTATGTAAGATCTGATTTGCTAGTTCTCC 1160  
 DB 177978 GTTTCACGCAATTAAGTTTATGAGATTTATGTAAGATCTGATTTGCTAGTTCTCC 178037  
 QY 1161 TTGTAGATTTAAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1220  
 DB 178038 TGTAGATTTAAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 178097  
 QY 1221 TATATTTGTGCTGCAAGATATTTGTAATTTGTTGTCACACTATGTAACAAACAGTGA 1280  
 DB 178098 TATATTTGTGCTGCAAGATATTTGTAATTTGTTGTCACACTATGTAACAAACAGTGA 178154  
 QY 1281 GATATGTTTAAATTAATTTGACTTATTTGGAAGTAAAAAAA 1323  
 DB 178155 GATATGTTTAAATTAATTTGACTTATTTGGAAGTAAAAAAA 178197

RESULT 10  
 AL512349 116614 bp DNA linear HTG 10-JUL-2001  
 LOCUS Homo sapiens chromosome 6 clone RP11-223J2, \*\*\* SEQUENCING IN  
 DEFINITION PROGRESS \*\*\*, in ordered pieces.  
 ACCESSION AL512349  
 VERSION AL512349.7 GI:13121473  
 KEYWORDS HTG; HTGS\_PHAZE2; HTGS\_CANCELLED.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 116614)  
 AUTHORS Bates, K.  
 JOURNAL Direct Submission  
 Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humbrey@sanger.ac.uk  
 On Feb 23, 2001 this sequence version replaced gi:1295681.  
 COMMENT ----- Genome Center  
 Center: Sanger Centre

Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk  
 Project Information  
 Center project name: B422302  
 ----- Summary Statistics -----  
 Assembly program: XGAP; Version 4.5  
 Sequencing vector: Plasmid; L08752; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Consensus quality: 115897 bases at least Q40  
 Consensus quality: 116315 bases at least Q30  
 Consensus quality: 116477 bases at least Q20  
 Insert size: 116614; sum-of-conligs  
 Insert coverage: 6.23x in Q20 bases; agarose-fp  
 coverage: 6.37x in Q20 bases; agarose-fp  
 -----  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 Location/Qualifiers  
 source  
 1. 116614  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="RP11-223j2"  
 /clone\_1lb="RP11-11.1"  
 misc\_feature  
 1. 116614  
 /note="assembly\_fragment:02232  
 clone\_end:SP6  
 vector\_slide:flight"  
 BASE COUNT 34471 a 21866 c 22518 g 37753 t 6 others  
 ORIGIN  
 Query Match 48.9% Score 687.6; DB 2; Length 116614;  
 Best Local Similarity 74.98; Pred. No. 1.6e-101; Indels 150; Gaps 13;  
 Matches 1099; Conservative 0; Mismatches 219;  
 QY 27 TGATGTGTTAGGAGGAGAGTAGGGAGAGCTTGTCTCCCGAGACGCTATCT 86  
 DB 10557 TGATGTGTTAGGAGGAGAGTAGGGAGAGCTTGTCTCTCAAGACGCTATCT 10498  
 QY 87 CATTCCTTCTTTCGATTACCGGTGGCGGGAGAGTCAGGGCGGGCGGTGGCGAGCAG 146  
 DB 10497 CATTCCTTCTTTCGATTACCGGTGGCGGGAGAGTCAGGGCGGGCGGTGGCGAGCAG 10467  
 QY 147 GCGGGGGGTGGCGGGCGGACCTGCATGATGTCAGCATGAAATCCGAATATGAT 206  
 DB 10466 AGTCATGGGTGGCGGAGCAGCTGCATGATGTCAGCATGAAATCCGAATATGAT 10407  
 QY 207 TATTATCAAGTACTTCTGATTGGCGACACAGGGGTGGAAAGTCCTTCTCTT 266  
 DB 10406 TATTATCAAGTACTTCTGATTGGCGACACAGGGGTGGAAAGTCCTTCTCTT 10347  
 QY 267 AGTTTGCAGATG-----ATACATATCAGAAAGCTACATCAGCAATTTGGGTG 317  
 DB 10346 AGTTTGCAGATG-----ATACATATCAGAAAGCTACATCAGCAATTTGGGTG 10287  
 QY 318 GATTCAAAATAGAGCTATAGAGTAGAGCGGAAACATCAAGCTTCAAT----- 371  
 DB 10286 AATTCAAAATAGAGCTATAGATATATATGCAAAATATCAAGCTTCAAGTGGAGC 10227  
 QY 372 -----GAGTCTTCATATATGTTAAACAGTGGCTGCAGAAATAGATGC 415  
 DB 10226 ACAACAGCGCAGAAATTTTATACCTCACTTATTACAGAGACCCAGGCTGTA 10167  
 QY 372 -----GAGTCTTCATATATGTTAAACAGTGGCTGCAGAAATAGATGC 415  
 DB 10166 TGATATGACAGATCAGAGGTCTTCGATATAGTTAAATATGTTGGTGCAGAAATGATCA 10107  
 QY 416 TTATGCACTGAAATGTCAACAATTTGGTATGAGACAAATGTATCTGACCACTAA 475  
 DB 10107 TTATGCACTGAAATGTCAACAATTTGGTATGAGACAAATGTATCTGACCACTAA 475

DB 10106 CTAGCCAGTGGAAATGTCAACAATTTGGTATGAGACAAATGTATCTGACCACTAA 10048  
 QY 476 GAAATAGTACTACACAAACAGCGAAGAAATTTGCTATTCCTTGGAAATTCCTTTT 535  
 DB 10047 GAA-----ACTAGTACTACACAAACAGCGAAGAAATTTGCTATTCCTTGGAAATTCCTTTT 9991  
 QY 536 GGAACAGAGTCTAGAAATGCAGAAATGTGAACAGTCTTTCATGACATGCGACGCA 595  
 DB 9990 GGAACAGAGTCTAGAAATGCAGAAATGTGAACAGTCTTTCATGAAATGCGATCGA 9931  
 QY 596 GATTAAAAAGGAATGGTCCCGAGCAACAAGCTGGTGTCTGAGAGTCAATGTAA 655  
 DB 9930 GATTAAAAAGGAATGGTCCCGAGCAACAAGCTGGTGTCTGAGAGTCAATGTAA 9872  
 QY 656 AATTAGAGCACTCCAGTCAAGCAGTACAGTGGAGGTGCTGTAAATTTCCCTCATC 715  
 DB 9871 AATTAGAGCACTCCAGTCAAGCAGTACAGTGGAGGTGCTGTAAATTTCCCTCATC 9812  
 QY 716 CTTTCTCAGCAATGAAATTTGCAATCTGATGACCAAGTGAATAAATTTGCTGAA 775  
 DB 9811 CTTTCTCAGCAATGAAATTTGCAATCTGATGACCAAGTGAATAAATTTGCTGAA 9753  
 QY 776 TTGTATGATAGTATGAGTGCAGTACACAGATTTCTTACGCTCCAGAAAGTCAAGAT 835  
 DB 9752 TTGTATGATAGTATGAGTGCAGTACACAGATTTCTTACGCTCCAGAAAGTCAAGAT 9693  
 QY 836 GAAATAGGCAATGAGTACTTTTATTTATCTCTTACCTGACCTCAACAGCTCAATCT 895  
 DB 9692 GAAATAGGCAATGAGTACTTTTATTTATCTCTTACCTGACCTCAACAGCTCAATCT 9634  
 QY 896 TCAGAACTGTTTAAACCTTGTGTGCTGTGTATTAATAATGTGTATCTCTTGTG 955  
 DB 9633 TCAGAACTGTTTAAACCTTGTGTGCTGTGTATTAATAATGTGTATCTCTTGTG 9574  
 QY 956 CTTTCTCAGTACCAAGCTTTCCCGTGTGGT-----AGAAATATTTTGTGTT 1007  
 DB 9573 CTTTCTCAGTACCAAGCTTTCCCGTGTGGT-----AGAAATATTTTGTGTT 9514  
 QY 1008 GATGTTTATATGCGATGTTTATGATGTCAGGTTTATGCTCTGAAGTGAAGTACGCA 1067  
 DB 9513 GATGTTTATATGCGATGTTTATGATGTCAGGTTTATGCTCTGAAGTGAAGTACGCA 9454  
 QY 1068 -TTTGTATCAAAACAGCAACAGAGTCTGTCTCCATTCATGATTAAGTTAAGTGA 1126  
 DB 9453 CTTTGTATCAAAACAGCAACAGAGTCTGTCTCCATTCATGATTAAGTTAAGTGA 9406  
 QY 1127 TGTATATGATATGATCTGATTTGCTAGTCTTCTCTGTAGAGTTATTAATGGAAGATTA 1186  
 DB 9405 TGTATATGATATGATCTGATTTGCTAGTCTTCTCTGTAGAGTTATTAATGGAAGATTA 9347  
 QY 1187 CACTATCTGATATATGTTCTTCTCATCTGATATTAATTTGGGTGCGAGATATTTG 1246  
 DB 9346 TACTATTTAATTAATGTTCTTCTCATCTGATATTAATTTGGGTGCGAGATATTTG 9287  
 QY 1247 AATTGTTGACACTGTGTACCAAAACATGAAATGTTTAAATATTTGACTTA 1306  
 DB 9286 GATTGTTGACACTGTGTACCAAAACATGAAATGTTTAAATATTTGACTTA 9228  
 QY 1307 TTGGAAGTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1366  
 DB 9227 CAATGATCTAATTTAAACCAAGAGCTTGTCTGAGAAATATTAACATGACAGAGGTG 9168  
 QY 1367 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1394  
 DB 9167 AATTAACACCAACCAAAATGAGAGAAA 9140  
 RESULT 11  
 HSDJ26307/c 104081 bp DNA linear PRI 23-NOV-1999  
 LOCUS  
 DEFINITION  
 Human DNA sequence from clone 26307 on chromosome 6q14.3-15.  
 Contains an RPL7 (60S Ribosomal protein L7) pseudogene, a RAB1  
 (RAB1, member RAS oncogene family) pseudogene, ESTs, an STS and

ACCESSION AL049545  
 VERSION HT: 605 Ribosomal Protein L7; RAB1; RAS; RPL7.  
 KEYWORDS Homo sapiens.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 104081)  
 AUTHORS Smith, S.  
 JOURNAL Direct Submission  
 Submitted (06-JUL-1999) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 Requests: clonerequests@sanger.ac.uk  
 On Jun 7, 1999 this sequence version replaced gi:4835284.  
 During sequence assembly data is compared from overlapping clones. While differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the validation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
 Em, EMBL; Sw, SWISSPROT; Tr, TREMBL  
 This sequence is the entire insert of clone 263J7. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the feature key.  
 This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>  
 263J7 is from the library RPL7 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:PCRPAC2>.  
 FEATURES  
 source  
 1..104081  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /map="q14.3-15"  
 /clone="RPL-263J7"  
 /clone\_id="RPL-1"  
 1..4880  
 /note="LIP5 repeat: matches -237. .5623 of consensus"  
 4895..5820  
 /note="LIP5 repeat: matches -1405. .-405 of consensus"  
 5861..6011  
 /note="LTR1 repeat: matches 36. .174 of consensus"  
 6032..6086  
 /note="L1 repeat: matches 3626. .3680 of consensus"  
 6740..8505  
 /note="LIP12 repeat: matches 4282. .6165 of consensus"  
 8512..9328  
 /note="LIP repeat: matches 3049. .3866 of consensus"  
 9329..9626  
 /note="AluYb8 repeat: matches 1. .317 of consensus"  
 9627..10966  
 /note="LIP repeat: matches 1712. .3049 of consensus"  
 /complement(11030..11749)  
 /gene="dJ263J7.1"  
 /gene="dJ263J7.1"  
 /complement(11030..11749)  
 /note="dJ263J7.1 (RPL7 (60S Ribosomal Protein L7) pseudogene); match: CDNAS: Em:X52567 Em:X57959 Em:U16558 Em:X57958 Em:M85235 Em:M17422 Em:M29016 Em:X57961 Em:X57960 Em:U63785; match: ESTs: Em:AL047212 Em:AI683841 Em:AI609411 Em:AI620251 Em:AI678713 Em:AA853290 Em:AI188448 Em:AI246632 Em:E29697.1 Em:AA640941 Em:AA147085 Em:AI354542 Em:AA68558 Em:AA172041

Em:AA736700 Em:AA181027 Em:AA617963 Em:AA129663  
 Em:AI385514 Em:AA586843 Em:AA148321 Em:AA147910  
 Em:AA065193 Em:AA223084 Em:AA055240 Em:AA314610  
 Em:AI498784 Em:AA313779 Em:AA626741 Em:AA397390  
 Em:AA639843 Em:AA598466 Em:AA567875 Em:AA314579  
 Em:AA081412 Em:AA838601 Em:AA129729 Em:AA236303  
 Em:AA580325 Em:AA147597 Em:AA313220 Em:AA516160  
 Em:AA311117 Em:AA562608 Em:AA635115 Em:AI367854  
 Em:AA308128 Em:AA846920 Em:AA608110 Em:AA745100  
 Em:AA607384 Em:AA507813 Em:AA314838 Em:AA736534  
 Em:AA591026 Em:AA008632 Em:W77465 Em:AA68073 Em:AA648202  
 Em:AA652160 Em:AA715370 Em:ID51697 Em:ID51979 Em:ID52053  
 Em:AA760175 Em:AA856854 Em:N98334 Em:AA649459 Em:AA879164  
 Em:AA332842 Em:AA888688 Em:AA173529 Em:AA091272  
 Em:AA093750; match: proteins: Sw:P18124 Sw:P14148  
 Sw:P05426 Sw:Q42208 Sw:O01802 Sw:Q60143 Sw:P25457  
 Sw:P11874 Sw:F32100 Sw:F05737 Sw:Q12213 Tr:O14731  
 Tr:O08953"  
 /pseudo  
 /codon\_start=1  
 /evidence-not\_experimental  
 11307..11789  
 /note="match: GSS: AQ081091 B54651 AQ028779"  
 11790..12860  
 /note="LIP13 repeat: matches 647. .1730 of consensus"  
 12908..13190  
 /note="Alu repeat: matches 3. .302 of consensus"  
 13341..13560  
 /note="HAI repeat: matches 1483. .1721 of consensus"  
 13346..13347  
 /note="3 subclones read TCTT, 2 read TCT"  
 13594..15880  
 /note="LIP2 repeat: matches 3950. .6154 of consensus"  
 15881..16608  
 /note="LIP4 repeat: matches 5413. .6144 of consensus"  
 16609..19090  
 /note="LIP2 repeat: matches 1394. .3950 of consensus"  
 19096..19323  
 /note="LIP3 repeat: matches 3. .231 of consensus"  
 19472..23244  
 /note="LIP repeat: matches 638. .4389 of consensus"  
 23240..23462  
 /note="LIPa repeat: matches -1549. .-1330 of consensus"  
 23577..23985  
 /note="LIP3 repeat: matches -218. .6130 of consensus"  
 29886..29943  
 /note="LIP5 repeat: matches 6081. .6139 of consensus"  
 29944..30234  
 /note="LIP3 repeat: matches -508. .-218 of consensus"  
 /complement(30558..30912)  
 /note="match: EST AA528504"  
 31385..31601  
 /note="MIR repeat: matches 29. .256 of consensus"  
 31721..32002  
 /note="LIP7 repeat: matches 5862. .6145 of consensus"  
 32279..32583  
 /note="AluX repeat: matches 1. .308 of consensus"  
 33356..33592  
 /note="MIR repeat: matches 8. .247 of consensus"  
 33779..37507  
 /note="LIM2 repeat: matches 2533. .6308 of consensus"  
 37508..37811  
 /note="AluY repeat: matches 3. .305 of consensus"  
 37812..37985  
 /note="LIM2 repeat: matches 2362. .2533 of consensus"  
 37976..41630  
 /note="LIM repeat: matches -1390. .2698 of consensus"  
 41878..41924  
 /note="MERB repeat: matches 127. .173 of consensus"  
 42325..42418  
 /note="MAD1 repeat: matches 1. .79 of consensus"  
 42451..42853  
 /note="LTR1 repeat: matches 23. .466 of consensus"



```

OY      1307 TTGCAAGTAAAAA... 1366
DB      94457 CAATGGATCTAATTAACCAAGAGCTTCGTCAGAAAATAACAGCAGAGCTG 94398
OY      1367 AAAAAAAAAAAAAAAAAAAAAA 1394
DB      94397 AATAGACACCAACCAAAATGAGAGAAA 94370

RESULT 12
AC067945/c 162471 bp DNA linear PRI 07-NOV-2001
LOCUS      AC067945 Homo sapiens BAC clone RP11-629B4 from 2, complete sequence.
DEFINITION AC067945
ACCESSION  AC067945.4 GI:13786486
VERSION     HTG.
KEYWORDS    Homo sapiens.
SOURCE      Homo sapiens.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 162471)
AUTHORS     Sulston,J.E. and Waterston,R.
TITLE       Toward a complete human genome sequence.
JOURNAL     Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE     99063792
PUBMED      9847074
REFERENCE   2 (bases 1 to 162471)
AUTHORS     Isak,A., Elliott,G., Doebber,A., Abbott,A., Hawkins,M. and Falk,A.
TITLE       The sequence of Homo sapiens BAC clone RP11-629B4
JOURNAL     Unpublished
REFERENCE   3 (bases 1 to 162471)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (27-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4 (bases 1 to 162471)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (25-APR-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5 (bases 1 to 162471)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (09-AUG-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Apr 25, 2001 this sequence version replaced g1:13518285.
COMMENT     ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: saplens@wustl.edu
            ----- Summary Statistics
            -----
            Center project name: H_NH0629B04

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence

from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateo,M., Catanesi,J.J. and de Jong,P.J. (1996) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

#### VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-45219, 200 bp overlap; the clone sequenced to the right is RP11-31739. Actual start of this clone is at base position 195 of RP11-629B4; actual end is at base position 162471 of RP11-629B4.

#### FEATURES

##### source

```

1..162471
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-629B4"
/clone_1lb="RPCI-11"
/rpt_family="AT_rich"
187..207
/rpt_family="MIR"
3209..3257
/rpt_family="MIR"
4158..4475
/rpt_family="L1"
4476..4775
/rpt_family="Alu"
4776..5044
/rpt_family="L1"
5108..5242
/rpt_family="Alu"
5428..5465
/rpt_family="AT_rich"
5617..5910
/rpt_family="Alu"
5997..6133
/rpt_family="Alu"
6135..6178
/rpt_family="Alu"
6213..6498
/rpt_family="Alu"
9226..9268
/rpt_family="AT_rich"
9573..9600
/rpt_family="AT_rich"
10528..10562
/rpt_family="AT_rich"
11440..11467
/rpt_family="(TTTATG)n"
12297..12325
/rpt_family="AT_rich"
12472..12518
/rpt_family="(TATATG)n"
14656..14747
/rpt_family="CRI"
16030..16089

```







DB 2459 CCTGTAGAGTATTAAGGACAGACACCTCTCTGATTAAGTTCTTCTACTCTG 2518  
|||||  
OY 1219 CATATATTCGTCGCGAGAAATATTCATTTGTCACACTCTGTACAAACAACTG 1278  
|||||  
DB 2519 CATATATTCGTCGCGAGAAATATTCATTTGTCACACTCTGTACAAACAACTG 2574  
|||||  
OY 1279 AAGATATGTTTAAATATTCATTTGTCACACTCTGTACAAACAACTG 1323  
|||||  
DB 2575 AAGATATGTTTAAATATTCATTTGTCACACTCTGTACAAACAACTG 2619  
|||||  
RESULT 14  
AC117841 207945 bp DNA linear HMG 17-JUL-2002  
LOCUS Rattus norvegicus clone CH230-360U6, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 73 unordered pieces.  
AC117841  
AC117841.4 GI:21746140  
HMG: HMGs.PHASE1.  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
AUTHORS 1 (bases 1 to 207945)  
Muzny D.M., Adams C., Adio-Oduola B., Ali-Osman F.R., Allen C.,  
Alsbrooks S.L., Amaratunga H.C., Are J.R., Ayele M., Banks T.,  
Barbata J., Benton J., Bimaga K., Blankenburg K., Bonini D.,  
Bouch J., Bowe S., Britva M., Brown E., Brown M., Bryant N.P.,  
Burch P., Burch P., Burkett C., Butrell K.L., Byrd N.C.,  
Carroll T.F., Carter M., Cavazos S.R., Chacko J., Chavaz D.,  
Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,  
Cleveland C.D., Cox C., Coyne M.D., Dathorne S.R., David R.,  
David M.L., Davis C., Davy-Carroll L., Dedrich D.A.,  
DeLaney R.R., Delgado O., Denn A.L., Ding X., Dinh H.H.,  
Douthett K.J., Draper H., Dugan-Rocha S., Durbin K.J.,  
Earhart C., Edgar D., Edwards C.C., Elhaj C., Escott M.,  
Falls T., Ferraguto D., Flagg N., Ford J., Foster P.,  
Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,  
Gorrell J.H., Guayra M., Gunaratne P., Hale S., Hamilton K.,  
Harris C., Harris K., Hart M., Hayak P., Hawes A., Hernandez J.,  
Hernandez O., Hodgson R., Hughes M., Holloway C., Hollins B.,  
Hommel F., Howard S., Huber J., Huynh S., Hume J., Jackson L.E.,  
Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,  
Karlsberg J., Kelly S., Khan U., King L., Koryak J., Kovar C.,  
Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C., Lewis L.,  
Li J., Li Z., Lichtarge O., Lieu C., Liu J., Liu W., Louissegh H.,  
Lozodo R.J., Lu X., Lucher A., Lucher R., Luna R., Ma J.,  
Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E.,  
Massey E., McInerney E., McLeod M.P., Meador M., Mei G., Melker M.,  
Miner G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S.,  
Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,  
Nguyen N., Nickerson E., Nwokenkwo S., Ogihara M., Ogihara G.,  
Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,  
Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,  
Rivers M., Rojka A., Rojokan I., Rolfe M., Ruiz S., Savery G.,  
Scherrer S., Scott G., Shen H., Shooshari N., Sisson L.,  
Sodergren E., Sonalke T., Sparks A., Stanley H., Stone H.,  
Sutton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,  
Tansey J., Taylor C., Taylor T., Teiford B., Thomas N., Thomas S.,  
Umanil K., Vasquez L., Vera Y., Villalon D., Vinson R., Wang O.,  
Wang S., Ward-Moore S., Warren R., Washington C., Wellington S.,  
Williams G., Williamson A., Wleczky R., Woodcock S., Worley K.,  
Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,  
Weinstock G., and Gibbs R.

REFERENCE  
AUTHORS Baylör Plaza, Houston, TX 77030, USA  
TITLE 3 (bases 1 to 207945)  
JOURNAL Direct Submission  
COMMENT Submitted (17-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:20258099.  
-----  
Center: Baylor College of Medicine  
Genome Center  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
-----  
Project Information  
Center project name: GRP  
Center clone name: CH230-360U6  
-----  
Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 129522 bases at least 940  
Consensus quality: 134968 bases at least 930  
Consensus quality: 138682 bases at least 920  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/centra\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/centra_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 73 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
1 1045: contig of 1046 bp in length  
\* 1047 1145: gap of unknown length  
\* 1147 2183: contig of 1037 bp in length  
\* 2184 2283: gap of unknown length  
\* 2284 3421: contig of 1138 bp in length  
\* 3422 3522: gap of unknown length  
\* 3522 4686: contig of 1165 bp in length  
\* 4687 4786: gap of unknown length  
\* 4787 6158: contig of 1372 bp in length  
\* 6159 6258: gap of unknown length  
\* 6259 7327: contig of 1069 bp in length  
\* 7328 7427: gap of unknown length  
\* 7428 9082: contig of 1655 bp in length  
\* 9083 9182: gap of unknown length  
\* 9183 10231: contig of 1049 bp in length  
\* 10232 10331: gap of unknown length  
\* 10332 11663: contig of 1332 bp in length  
\* 11664 13476: gap of unknown length  
\* 13477 13576: contig of 1713 bp in length  
\* 13577 15050: contig of 1474 bp in length  
\* 15051 15150: gap of unknown length  
\* 15151 16522: contig of 1372 bp in length  
\* 16523 16623: gap of unknown length  
\* 16624 18019: contig of 1397 bp in length  
\* 18020 18119: gap of unknown length  
\* 18120 19774: contig of 1655 bp in length  
\* 19775 19874: gap of unknown length  
\* 19875 21203: contig of 1329 bp in length  
\* 21204 22991: gap of unknown length  
\* 22992 23091: contig of 1688 bp in length  
\* 23092 24721: gap of unknown length  
\* 24722 24821: contig of 1630 bp in length  
\* 24822 25833: gap of unknown length  
\* 25834 27568: contig of 1012 bp in length  
\* 27569 27668: gap of unknown length  
\* 27669 29226: contig of 1635 bp in length  
\* 29227 29226: contig of 1558 bp in length

Query Match	Similarity	Score	DB 2:	length	207945:
Matches 937:	Conservative	0:	Mismatches 98:	Indels 109:	Gaps 9:
29327	29326	gap of unknown length			
29327	30423	contig of 1098 bp in length			
30425	30524	gap of unknown length			
30525	32145	contig of 1621 bp in length			
32146	32245	gap of unknown length			
32246	33498	contig of 1253 bp in length			
33499	33598	gap of unknown length			
33599	34724	contig of 1126 bp in length			
34725	34824	gap of unknown length			
34825	36502	contig of 1678 bp in length			
36503	36601	gap of unknown length			
36603	37971	contig of 1369 bp in length			
37972	38071	gap of unknown length			
38072	39263	contig of 1192 bp in length			
39264	39363	gap of unknown length			
39364	41128	contig of 1765 bp in length			
41129	41228	gap of unknown length			
41229	42650	contig of 1422 bp in length			
42651	42750	gap of unknown length			
42751	44150	contig of 1400 bp in length			
44151	44250	gap of unknown length			
44251	45440	contig of 1190 bp in length			
45441	45540	gap of unknown length			
45541	47375	contig of 1835 bp in length			
47376	47475	gap of unknown length			
47476	48634	contig of 1159 bp in length			
48635	48734	gap of unknown length			
48735	50080	contig of 1346 bp in length			
50081	50180	gap of unknown length			
50181	51471	contig of 1291 bp in length			
51472	51571	gap of unknown length			
51572	52959	contig of 1388 bp in length			
52960	53059	gap of unknown length			
53060	54842	contig of 1783 bp in length			
54843	54942	gap of unknown length			
54943	57384	contig of 2442 bp in length			
57385	57484	gap of unknown length			
57485	59403	contig of 1919 bp in length			
59404	59503	gap of unknown length			
59504	60963	contig of 1460 bp in length			
60964	61063	gap of unknown length			
61064	62330	contig of 1267 bp in length			
62331	62430	gap of unknown length			
62431	63700	contig of 1270 bp in length			
63701	63800	gap of unknown length			
63801	65110	contig of 1310 bp in length			
65111	66210	gap of unknown length			
66211	66895	contig of 1685 bp in length			
66896	66995	gap of unknown length			
66996	68391	contig of 1396 bp in length			
68392	68491	gap of unknown length			
68492	70672	contig of 2181 bp in length			
70673	70772	gap of unknown length			
70773	72373	contig of 1601 bp in length			
72374	72473	gap of unknown length			
72474	73328	contig of 2855 bp in length			
73329	75428	gap of unknown length			
75429	77871	contig of 2443 bp in length			
77872	77971	gap of unknown length			
77972	80743	contig of 2772 bp in length			
80744	80843	gap of unknown length			
80844	82486	contig of 1643 bp in length			
82487	82586	gap of unknown length			
82587	84732	contig of 2146 bp in length			
84733	84832	gap of unknown length			
84833	87219	contig of 2387 bp in length			
87220	87319	gap of unknown length			
87320	90363	contig of 3044 bp in length			

Db	90783	GGGCTTTTTCGAATATTTCTTTTGATGTTATATGGCATCTTACATCTGGGT	90724
OY	1041	TAGT	1044
Db	90723	TAGT	90720



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 18:55:48 ; Search time 351 Seconds  
(without alignments)  
9014.409 Million cell updates/sec

Title: US-09-820-003b-1

Perfect score: 1405  
Sequence: 1 aagcagatgctgagtcgagc.....aaaaaaaaaaaaaaaaaaaa 1405

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200.2	85.4	2528	21	AA600009 Human secreted pro
2	1029	73.2	1193	23	AA583860 cDNA encoding novel
3	506	36.0	506	24	ABR44823 Human cDNA diffe
4	476.2	33.9	723	24	ABR83978 Human breast cell
5	467	33.2	487	22	ABN43096 Human foetal liver
6	467	33.2	487	22	ABN53512 Probe #1742 for ge
7	467	33.2	487	22	ABN23276 Human brain expres
8	467	33.2	487	22	AAK01783 Human bone marrow
9	467	33.2	487	22	AAK27239

C	10	467	33.2	487	22	AA111818	Probe #1751 for ge
C	11	467	33.2	487	22	AA133136	Probe #1822 used t
C	12	467	33.2	487	22	AA101752	Probe #1743 used t
C	13	467	33.2	487	24	AB501772	Human genome-deriv
C	14	333.4	23.7	335	21	AA606621	Human secreted pro
C	15	264	18.8	406	23	ABV14061	Human prostate exp
C	16	264	18.6	448	23	ABV35165	Human prostate exp
C	17	262	18.6	448	23	ABV44007	Human prostate exp
C	18	261	18.6	200	24	ABV45210	Corn tassal-derive
C	19	232.8	16.6	241	22	AA547279	Human breast cancer
C	20	232.8	16.6	241	22	AA547279	Human breast cancer
C	21	222.4	15.8	250	20	AAV04892	Human prostate exp
C	22	222.4	15.8	250	20	AAV04892	Human prostate exp
C	23	219.8	15.6	229	22	AAH55559	EST clone FC327
C	24	195	13.9	925	20	AAH27232	Human breast tumor
C	25	195	13.9	939	21	AAH21661	Human Rab protein,
C	26	195	13.9	1898	22	AA544924	Human breast and o
C	27	195	13.9	1944	22	AA544924	Human polynucleoti
C	28	190.8	13.6	737	23	AA583862	DNA encoding novel
C	29	188.8	13.4	214	21	AA542979	Human secreted exp
C	30	187	13.3	1965	22	AA534668	Human DNM for a no
C	31	179	12.7	179	22	ABA48222	Human secreted cell
C	32	179	12.7	179	22	ABA66102	Human foetal liver
C	33	179	12.7	179	22	ABA33177	Probe #11643 for g
C	34	179	12.7	179	22	AAK14525	Human brain expres
C	35	179	12.7	179	22	AAK40268	Human bone marrow
C	36	179	12.7	179	22	AAI21030	Probe #1093 for g
C	37	179	12.7	179	22	AAI46285	Probe #1937 used f
C	38	179	12.7	179	22	AAI06747	Human genome-deriv
C	39	179	12.7	179	22	AAI4239	Human gene signal c
C	40	175.2	12.5	165	16	AAI19274	Human neutrophil c
C	41	171	11.8	171	20	AAK23507	Human neutrophil c
C	42	163	11.6	166	20	AAK23476	Drosophila melanog
C	43	163	11.6	1791	23	ABL05963	Peppermint plant o
C	44	155.4	11.1	607	22	AAH87926	Arabidopsis thalia
C	45	150.8	10.7	932	21	AA635200	

#### ALIGNMENTS

RESULT 1	AA600009	standard; cDNA; 2528 BP.
ID	AA600009	
XX	AA600009	
XX	26-JAN-2001	(first entry)
XX	Human secreted protein gene 44 SEQ ID NO:34.	
XX	Human: secreted protein; diagnosis; cytosolic; immunosuppressive;	
XX	nocturnal; neuroprotective; antiviral; antiallergic; hepatotropic;	
XX	antidiabetic; antiinflammatory; antitumor; anticonvulsant;	
XX	antibacterial; antifungal; antiparasitic; cardiant; gene therapy;	
XX	cancer; immune disorder; cardiovascular disorder; wound healing;	
XX	neurological disease; infectious disease; chromosome identification; ss.	
OS	Homo sapiens.	
XX	WO200058356-A1.	
XX	05-OCT-2000.	
XX	22-MAR-2000; 2000WO-US07535.	
XX	26-MAR-1999; 99US-0126511.	
XX	17-DEC-1999; 99US-0172413.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
XX	Rosen CA, Ruben SM, Komatsoulis G;	
XX		



Human: chromosome mapping; gene mapping; gene therapy; forensic: food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX MO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001MO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dermanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; AB019673.

XX New isolated polynucleotide and encoded polypeptides, useful in PT diagnostics, forensics, gene mapping, identification of mutations PT responsible for genetic disorders or other traits and to assess PT biodiversity.

XX Claim 1; SEQ ID No 19664; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and PS polypeptide (II) sequences. (I) is useful as hybridisation probes, CC polypeptide chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (I) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving CC quantifying a polypeptide in tissue, as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. CC The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64197-AAS94564 represent novel human CC diagnostic coding sequences of the invention. CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC atftp.wipo.int/pub/published\_pcr\_sequences.

XX Sequence 1193 BP; 372 A; 279 C; 213 G; 329 T; 0 other;

XX Query Match 73.2%; Score 1029; DB 23; Length 1193;

XX Best Local Similarity 89.9%; Pred. No. 8-7e-132;

XX Matches 1182; Conservative 0; Mismatches 0; Indels 133; Gaps 2;

QY 1 AAGGATAGCTAGT-GGGGCGCTGCTGATTTGTTCTAGGGAGAGTATGGGGAAGA 59

DB 1183 AAGGATAGCTAGTGGGGGCGGCTGCTGATTTGTTCTAGGGAGAGTATGGGGAAGA 1124

QY 60 GCTTTCCTCCCGGAGACGCTATCTCATTTCTTTTCGATTAACCGTGGCGCGAG 119

DB 1123 GCTTTCCTCCCGGAGACGCTATCTCATTTCTTTTCGATTAACCGTGGCGCGAG 1064

QY 120 AGTCAGGCGGCGCTGGCGAGCAAGGCGGCGGTTGGCGGCGGCGAGCTGAGTGAC 179

DB 1063 AGTCAGGCGGCGGCTGGCGAGCAAGGCGGCGGTTGGCGGCGGCGAGCTGAGTGAC 1004

QY 180 ATGTCCAGCATGATCCGATATGATATTTATTTATTTCAAGTACTTCTGATGGGACTCA 239

DB 1003 ATGTCCAGCATGATCCGATATGATATTTATTTATTTCAAGTACTTCTGATGGGACTCA 944

QY 240 GGGGTGGAAGCTTGCTCTCTTAGGTTTGAGATGATATATACAGAAAGCTAC 299

DB 943 GGGGTGGAAGCTTGCTCTCTTAGGTTTGAGATGATATATACAGAAAGCTAC 884

QY 300 ATACACAAATTTGGTGTGATTTTCAAAATAGAACTATAGAGTTAGACGGGAAACATC 359

DB 883 ATACACAAATTTGGTGTGATTTTCAAAATAGAACTATAGAGTTAGACGGGAAACATC 824

QY 360 AACCTTCAATTTAGATCTTTCATATATGTTTAAACAGTGGCTGCAGAAATAGATGTTAT 419

DB 823 AACCTTCAATTTAGATCTTTCATATATGTTTAAACAGTGGCTGCAGAAATAGATGTTAT 812

QY 420 GCCAGTGAATTTAGTCAACAATTTGGTGTAGGGAACAATGTGATCTGACACAAAGAA 479

DB 811 ----- 812

QY 480 GTATAGACTACACACAGGAAGAAATTTGCTGATTCCTTGGATTCGTTTTGGAA 539

DB 811 -----GAATTTGCTGATTCCTTGGATTCGTTTTGGAA 776

QY 540 ACCAGTGTAAAGATGCAAGCAATGTAGACAGTCTTTCATGACGATGGACCTGATTT 599

DB 775 ACCAGTGTAAAGATGCAAGCAATGTAGACAGTCTTTCATGACGATGGACCTGATTT 716

QY 600 AAAAAGCAATGGGTCCGGAGCAACAGCTGGTGGCTGAGAGTCCAAATGTTAAAT 659

DB 715 AAAAAGCAATGGGTCCGGAGCAACAGCTGGTGGCTGAGAGTCCAAATGTTAAAT 656

QY 660 CAGAGCTCTCAGTCAAGCAAGTCAAGTGGAGAGTGGCTGTAATTTGCTGCATCTTT 719

DB 655 CAGAGCTCTCAGTCAAGCAAGTCAAGTGGAGAGTGGCTGTAATTTGCTGCATCTTT 596

QY 720 TCTCAGACATGATTTGCAATCTGAAACCCAGTGAACCAAAATTTGCTGATTT 779

DB 595 TCTCAGACATGATTTGCAATCTGAAACCCAGTGAACCAAAATTTGCTGATTT 536

QY 780 ACTGATGATGATGCTACATCAACAGATTTTACCTGCTCCACAAAGTCAAGATTTAA 839

DB 535 ACTGATGATGATGCTACATCAACAGATTTTACCTGCTCCACAAAGTCAAGATTTAA 476

QY 840 ATGTCCATATGATCTTTTATTTTATTCCTTGACTCAACAGCTTAATCTTATTTGAG 899

DB 475 ATGTCCATATGATCTTTTATTTTATTCCTTGACTCAACAGCTTAATCTTATTTGAG 416

QY 900 AACTGTTTAAACCTTTTGCTGCTGCTGTTTAAATATATGCTGTAATTCCTGCTGCTT 959

DB 415 AACTGTTTAAACCTTTTGCTGCTGCTGTTTAAATATATGCTGTAATTCCTGCTGCTT 356

QY 960 CCGATACACAGACTGTTTCCGCTGGTGGTTAGCAATATATTTGTTGATGTTTATTT 1019

DB 355 CCGATACACAGACTGTTTCCGCTGGTGGTTAGCAATATATTTGTTGATGTTTATTT 296

QY 1020 GGCATGTTTATGATGCTAGGTTTATGCTTCTGAAAGATGAAGTTCAGCCATTTGTTCAAA 1079

DB 295 GGCATGTTTATGATGCTAGGTTTATGCTTCTGAAAGATGAAGTTCAGCCATTTGTTCAAA 236

QY 1080 CAGCACAGCAGTGTCTGTCACATTTCCATGATTAAGTTTATGATGATGTTTATGTAAG 1139

DB 235 CAGCACAGCAGTGTCTGTCACATTTCCATGATTAAGTTTATGATGATGTTTATGTAAG 176

QY 1140 ATCTGATTTGCTAGTCTTCTCTTGTAGAGTATTAATGAAAGATTTACACTATTCGATTA 1199

DB 175 ATCTGATTTGCTAGTCTTCTCTTGTAGAGTATTAATGAAAGATTTACACTATTCGATTA 116

QY 1200 ATAGTTTCTTATACCTGCTCATATATTTTGGCTGCAAGATATTTGTAATTTGGTCA 1259

DB 115 ATAGTTTCTTATACCTGCTCATATATTTTGGCTGCAAGATATTTGTAATTTGGTCA 56

QY 1260 CTATGTAAACAAACAACTGAAGATATGTTTAAATTAATTTTGTACTTTTGGAGAT 1314

DB 55 CTATGTAAACAAACAACTGAAGATATGTTTAAATTAATTTTGTACTTTTGGAGAT 1

RESULT 3







SO Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match 33.2%; Score 467; DB 22; Length 487;

Best Local Similarity 100.0%; Pred. No. 2.4e-55;

Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 502 AGAATTTGCGATTCCTTGGAAATTCGGTTTTGGAAACAGTCTAAGATGCAACGA 561
DB 467 AGGATTTGCGATTCCTTGGAAATTCGGTTTTGGAAACAGTCTAAGATGCAACGA 408
OY 562 ATGTAGAAGACGCTTTCATGAGAGTGGAGTGGATTTAAAGAGATGGTCCCGAG 621
DB 407 ATGTAGAAGACGCTTTCATGAGAGTGGAGTGGATTTAAAGAGATGGTCCCGAG 348
OY 622 CAACAGCTGTGTGCTGAGAGTCCATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 681
DB 347 CAACAGCTGTGTGCTGAGAGTCCATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 288
OY 682 CAGGTGAGGTGCTGCTAAATTTGCTCCATCCCTTTTCACAGCAATGAATTTGCA 741
DB 287 CAGGTGAGGTGCTGCTAAATTTGCTCCATCCCTTTTCACAGCAATGAATTTGCA 228
OY 742 TCTGACCCCAAGTGAACAAATTCCTGAATTTGCTGATGATGCTGCACTACAA 801
DB 227 TCTGACCCCAAGTGAACAAATTTGCTGAATTTGCTGATGATGCTGCACTACAA 168
OY 802 CAGATTTCTACCGTCTCCACAAGGTCAAGATTTGTAATGGTCAATGACTGCTTTT 861
DB 167 CAGATTTCTACCGTCTCCACAAGGTCAAGATTTGTAATGGTCAATGACTGCTTTT 108
OY 862 TTATTCCTTACGACCAACAGCTCACTTCAATTTTCAACAGTCTTTAACTTTGTG 921
DB 107 TTATTCCTTACGACCAACAGCTCACTTCAATTTTCAACAGTCTTTAACTTTGTG 48
OY 922 CTGCTTTTAAATTAATGTGTGTAATCCTTGTGCTTCCGTATACC 968
DB 47 CTGCTTTTAAATTAATGTGTGTAATCCTTGTGCTTCCGTATACC 1

```

#### RESULT 6

ABAS3512/c

ID ABA3512 standard; DNA; 487 BP.

XX ABA3512;

XX 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #1817.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX

PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human fetal liver -  
 PS Claim 1; SEQ ID NO 1817; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC foetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pcr\_sequences.

SO Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match 33.2%; Score 467; DB 22; Length 487;

Best Local Similarity 100.0%; Pred. No. 2.4e-55;

Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 502 AGAATTTGCGATTCCTTGGAAATTCGGTTTTGGAAACAGTCTAAGATGCAACGA 561
DB 467 AGGATTTGCGATTCCTTGGAAATTCGGTTTTGGAAACAGTCTAAGATGCAACGA 408
OY 562 ATGTAGAAGACGCTTTCATGAGAGTGGAGTGGATTTAAAGAGATGGTCCCGAG 621
DB 407 ATGTAGAAGACGCTTTCATGAGAGTGGAGTGGATTTAAAGAGATGGTCCCGAG 348
OY 622 CAACAGCTGTGTGCTGAGAGTCCATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 681
DB 347 CAACAGCTGTGTGCTGAGAGTCCATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 288
OY 682 CAGGTGAGGTGCTGCTAAATTTGCTCCATCCCTTTTCACAGCAATGAATTTGCA 741
DB 287 CAGGTGAGGTGCTGCTAAATTTGCTCCATCCCTTTTCACAGCAATGAATTTGCA 228
OY 742 TCTGACCCCAAGTGAACAAATTCCTGAATTTGCTGATGATGCTGCACTACAA 801
DB 227 TCTGACCCCAAGTGAACAAATTCCTGAATTTGCTGATGATGCTGCACTACAA 168
OY 802 CAGATTTCTACCGTCTCCACAAGGTCAAGATTTGTAATGGTCAATGACTGCTTTT 861
DB 167 CAGATTTCTACCGTCTCCACAAGGTCAAGATTTGTAATGGTCAATGACTGCTTTT 108
OY 862 TTATTCCTTACGACCAACAGCTCACTTCAATTTTCAACAGTCTTTAACTTTGTG 921
DB 107 TTATTCCTTACGACCAACAGCTCACTTCAATTTTCAACAGTCTTTAACTTTGTG 48
OY 922 CTGCTTTTAAATTAATGTGTGTAATCCTTGTGCTTCCGTATACC 968
DB 47 CTGCTTTTAAATTAATGTGTGTAATCCTTGTGCTTCCGTATACC 1

```

#### RESULT 7

ABA23276/c

ID ABA23276 standard; DNA; 487 BP.

XX ABA23276;

XX 23-JAN-2002 (first entry)

DE Probe #1742 for gene expression analysis in human heart cell sample.

KW Human; gene expression; heart; microarray; vascular system; probe;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease; ss.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX



Db 227 TCTAACCAGTAAACAAAAATGCTGAAATGCTGTAATGTAAGTGCACATACAA 168  
 Oy 802 CAGATTCCTACCGTCTCCACAAAGGTGAGATTTGTAATGATCAATACCTATTTT 861  
 Db 167 CAGATTCCTACCGTCTCCACAAAGGTGAGATTTGTAATGATCAATACCTATTTT 108  
 Oy 862 TTATTCCTGAGTCAAGACAGCTAATCTTCAATTTTCAGAACTGTTTAAACCTTGTG 921  
 Db 107 TTATTCCTGAGTCAAGACAGCTAATCTTCAATTTTCAGAACTGTTTAAACCTTGTG 48  
 Oy 922 CTGTTTAAATAATGATGCTGTAATGCTGCTGCTGCTGATACC 968  
 Db 47 CTGTTTAAATAATGATGCTGTAATGCTGCTGCTGCTGATACC 1

RESULT 9  
 AAK27239/C  
 ID AAK27239 standard; DNA; 487 BP.

AAK27239;

06-NOV-2001 (first entry)

Human bone marrow expressed single exon probe SEQ ID NO: 1796.

Human; bone marrow expressed exon; gene expression analysis; probe; microarray; cancer; leukemia; lymphoma; myeloma; ss.

Homo sapiens.

WO200157276-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00668.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human bone marrow -

Example 4; SEQ ID NO: 1796; 658bp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukemia and myeloma. The present sequence is one of the probes of the invention.

Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match 33.2%; Score 467; DB 22; Length 487;

Best Local Similarity 100.0%; Pred. No. 2.4e-55; Mismatches 0; Indels 0; Gaps 0;

502 AGGAATTTGCTGATTCCTTGGAAATTCGTTTGGAAACGAGTGAAGATGCAACA 561  
 467 AGGAATTTGCTGATTCCTTGGAAATTCGTTTGGAAACGAGTGAAGATGCAACA 408

Oy 562 ATGTAGAACAGCTCTTTCATGACGATGGCAGCTGAGATTAACAAAGCAATGGTCCGAG 621  
 Db 407 ATGTAGAACAGCTCTTTCATGACGATGGCAGCTGAGATTAACAAAGCAATGGTCCGAG 348  
 Oy 622 CAACAGCTGGTGGTGTGAGAGATTCGAATGTTAAATTCAGAGATTCAGTCAAGAGT 681  
 Db 347 CAACAGCTGGTGGTGTGAGAGATTCGAATGTTAAATTCAGAGATTCAGTCAAGAGT 288  
 Oy 682 CAGGTGAGAGTGTGCTGTAATAATTTGCCATCCTTTTCTCAGAGCAATGATTTGCA 741  
 Db 287 CAGGTGAGAGTGTGCTGTAATAATTTGCCATCCTTTTCTCAGAGCAATGATTTGCA 228  
 Oy 742 TCTGACCCAAAGTAAACAAATTCCTGAAATTTGCTGATGATGCTGACATCAAA 801  
 Db 227 TCTGACCCAAAGTAAACAAATTCCTGAAATTTGCTGATGATGCTGACATCAAA 168  
 Oy 802 CAGATTCCTACCGTCTCCACAAAGGTGAGATTTGTAATGATCAATACCTATTTT 861  
 Db 167 CAGATTCCTACCGTCTCCACAAAGGTGAGATTTGTAATGATCAATACCTATTTT 108  
 Oy 862 TTATTCCTGAGTCAAGACAGCTAATCTTCAATTTTCAGAACTGTTTAAACCTTGTG 921  
 Db 107 TTATTCCTGAGTCAAGACAGCTAATCTTCAATTTTCAGAACTGTTTAAACCTTGTG 48  
 Oy 922 CTGTTTAAATAATGATGCTGTAATGCTGCTGCTGCTGATACC 968  
 Db 47 CTGTTTAAATAATGATGCTGTAATGCTGCTGCTGCTGATACC 1

RESULT 10  
 AAl1818/C  
 ID AAl1818 standard; DNA; 487 BP.

AAl1818;

12-OCT-2001 (first entry)

Probe #1751 for gene expression analysis in human cervical cell sample.

Probe; human; microarray; gene expression; cervical epithelial cell;

cervical cancer; ss.

Homo sapiens.

WO200157278-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00670.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488901/53.

Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human cervical epithelial cells -

Claim 25; SEQ ID NO 1751; 487bp; English.

The present invention relates to human single exon nucleic acid probes (SENPs). The present sequence is one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a

CC sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer.

CC Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pcf\\_sequences](http://wipo.int/pub/published_pcf_sequences).

XX Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other:

SO

Query Match 33.28; Score 467; DB 22; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-55;  
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 502 AGGAATTTGGTGAATCCCTTGAATTCCTTTTGGAAACCAAGTCGTAAGTAAGCAAGA 561  
 |||||||  
 DB 467 AGGAATTTGGTGAATCCCTTGAATTCCTTTTGGAAACCAAGTCGTAAGTAAGCAAGA 408  
 |||||||  
 OY 562 ATGTAGAACAGTCTTTCATGACGATGCGACGTGAGATTAAACCAATGGTCCCGAG 621  
 |||||||  
 DB 407 ATGTAGAACAGTCTTTCATGACGATGCGACGTGAGATTAAACCAATGGTCCCGAG 348  
 |||||||  
 OY 622 CACACAGCTGGTGGTGGTGGAGAGTCACATGTGTAATTCACAGCAGCTCCACAGCAGT 681  
 |||||||  
 DB 347 CACACAGCTGGTGGTGGTGGAGAGTCACATGTGTAATTCACAGCAGCTCCACAGCAGT 288  
 |||||||  
 OY 682 CAGGTGGAGGTGCTCTCTTAAATTTGGCTCAGTCCCTTTTCACAGCAGTAAGTTGCA 741  
 |||||||  
 DB 287 CAGGTGGAGGTGCTCTCTTAAATTTGGCTCAGTCCCTTTTCACAGCAGTAAGTTGCA 228  
 |||||||  
 OY 742 TCTGACCCCAAGTCAAAAAAGAAATGGCTGATGTGATGATGATGATGATGATGATGAT 801  
 |||||||  
 DB 227 TCTGACCCCAAGTCAAAAAAGAAATGGCTGATGTGATGATGATGATGATGATGATGAT 168  
 |||||||  
 OY 802 CAGATCTTACGCTCCCAAGAGTCAGAGATGTAATTAATGTCATGACTGCTTTT 861  
 |||||||  
 DB 167 CAGATCTTACGCTCCCAAGAGTCAGAGATGTAATTAATGTCATGACTGCTTTT 108  
 |||||||  
 OY 862 TTAATCCCTGACTCAAGACGTAAGTCAATTCATGCTTCAAGAGTGTAAACCTTGTGTG 921  
 |||||||  
 DB 107 TTAATCCCTGACTCAAGACGTAAGTCAATTCATGCTTCAAGAGTGTAAACCTTGTGTG 48  
 |||||||  
 OY 922 CTGGTTTAAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 968  
 |||||||  
 DB 47 CTGGTTTAAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 11  
 AA13136/c  
 ID AA13136 standard; DNA; 487 BP.  
 XX  
 AC AA13136;  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Probe #1822 used to measure gene expression in human placenta sample.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder; s8.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001MO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236559.

PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 PS Claim 25; SEQ ID No 1822; 654bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SNP).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.

SO Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other:

Query Match 33.28; Score 467; DB 22; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-55;  
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 502 AGGAATTTGGTGAATCCCTTGAATTCCTTTTGGAAACCAAGTCGTAAGTAAGCAAGA 561  
 |||||||  
 DB 467 AGGAATTTGGTGAATCCCTTGAATTCCTTTTGGAAACCAAGTCGTAAGTAAGCAAGA 408  
 |||||||  
 OY 562 ATGTAGAACAGTCTTTCATGACGATGCGACGTGAGATTAAACCAATGGTCCCGAG 621  
 |||||||  
 DB 407 ATGTAGAACAGTCTTTCATGACGATGCGACGTGAGATTAAACCAATGGTCCCGAG 348  
 |||||||  
 OY 622 CACACAGCTGGTGGTGGTGGAGAGTCACATGTGTAATTCACAGCAGCTCCACAGCAGT 681  
 |||||||  
 DB 347 CACACAGCTGGTGGTGGTGGAGAGTCACATGTGTAATTCACAGCAGCTCCACAGCAGT 288  
 |||||||  
 OY 682 CAGGTGGAGGTGCTCTCTTAAATTTGGCTCAGTCCCTTTTCACAGCAGTAAGTTGCA 741  
 |||||||  
 DB 287 CAGGTGGAGGTGCTCTCTTAAATTTGGCTCAGTCCCTTTTCACAGCAGTAAGTTGCA 228  
 |||||||  
 OY 742 TCTGACCCCAAGTCAAAAAAGAAATGGCTGATGTGATGATGATGATGATGATGATGAT 801  
 |||||||  
 DB 227 TCTGACCCCAAGTCAAAAAAGAAATGGCTGATGTGATGATGATGATGATGATGATGAT 168  
 |||||||  
 OY 802 CAGATCTTACGCTCCCAAGAGTCAGAGATGTAATTAATGTCATGACTGCTTTT 861  
 |||||||  
 DB 167 CAGATCTTACGCTCCCAAGAGTCAGAGATGTAATTAATGTCATGACTGCTTTT 108  
 |||||||  
 OY 862 TTAATCCCTGACTCAAGACGTAAGTCAATTCATGCTTCAAGAGTGTAAACCTTGTGTG 921  
 |||||||  
 DB 107 TTAATCCCTGACTCAAGACGTAAGTCAATTCATGCTTCAAGAGTGTAAACCTTGTGTG 48  
 |||||||  
 OY 922 CTGGTTTAAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 968  
 |||||||  
 DB 47 CTGGTTTAAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 12  
 AA101752/c  
 ID AA101752 standard; DNA; 487 BP.  
 XX  
 AC AA101752;  
 DT 09-OCT-2001 (first entry)  
 XX  
 DE Probe #1743 used to measure gene expression in human breast sample.  
 XX  
 KW Probe; human; breast disease; breast cancer; development disorder; s8;  
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 XX  
 OS Homo sapiens.  
 XX





CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences; mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene  
CC expression analysis, and for identifying exons in a gene, particularly  
CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic  
CC and hyaline membrane disease. The present sequence is a single exon  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIP0 at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match 33.2%; Score 467; DB 24; Length 487;

Best Local Similarity 100.0%; Pred. No. 2.4e-55; Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 502 AGGAATTTGCTGATTCCTTGAATTCCTTTTGAACAGAGGTAGAAATGCAACA 561  
|||||  
DB 467 AGGAATTTGCTGATTCCTTGAATTCCTTTTGAACAGAGGTAGAAATGCAACA 408  
|||||  
DB 562 ATGAGAACAGCTTTTCATGACGATGAGAGCTGATTAACAGATGGGTCCGGAG 621  
|||||  
DB 407 ATGAGAACAGCTTTTCATGACGATGAGAGCTGATTAACAGATGGGTCCGGAG 348  
|||||  
DB 622 CAACAGCTGGTGGTCTGAGAGCTGATTAACAGATGGGTCCGGAG 661  
|||||  
DB 347 CAACAGCTGGTGGTCTGAGAGCTGATTAACAGATGGGTCCGGAG 288  
|||||  
DB 682 CAGGTGAGAGCTGGTCTGAGAGCTGATTAACAGATGGGTCCGGAG 741  
|||||  
DB 287 CAGGTGAGAGCTGGTCTGAGAGCTGATTAACAGATGGGTCCGGAG 228  
|||||  
DB 742 TCTGACCCAGTGAACCAAAATTCCTGATTTAGTGTAGCTGACACTACAA 801  
|||||  
DB 227 TCTGACCCAGTGAACCAAAATTCCTGATTTAGTGTAGCTGACACTACAA 168  
|||||  
DB 802 CAGATTTTACGCTGTCACCAAGGTCAGATTTGTAATGGTCAATGACTTTT 861  
|||||  
DB 167 CAGATTTTACGCTGTCACCAAGGTCAGATTTGTAATGGTCAATGACTTTT 108  
|||||  
DB 862 TTAATCCCTGACTGACAGACAGTAACTTATTCAGACTGTTTAAACCTTTGTC 921  
|||||  
DB 107 TTAATCCCTGACTGACAGACAGTAACTTATTCAGACTGTTTAAACCTTTGTC 48  
|||||  
DB 922 CTGGTTATTAATAATGTTGTAATGCTTTGCTTCCGTGATAC 968  
|||||  
DB 47 CTGGTTATTAATAATGTTGTAATGCTTTGCTTCCGTGATAC 1  
|||||

## RESULT 14

AA06621 standard; cDNA: 335 BP.

XX AAC06621;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 10696.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.  
XX 06-SEP-2000.  
XX 21-FEB-2000; 2000EP-0200610.  
XX 26-FEB-1999; 9905-0122487.  
XX (GIST ) GENSET.  
XX Dunas Milne Edwards J, Duclert A, Giordano J;  
XX WPI: 2000-500381/45.  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
XX Claim 1; SEQ ID 10696; 71bp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from  
CC cDNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or poly(A) RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors.

XX Sequence 335 BP; 96 A; 70 C; 56 G; 109 T; 4 other;

Query Match 23.7%; Score 333.4; DB 21; Length 335;

Best Local Similarity 98.8%; Pred. No. 3.6e-37; Matches 331; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

DB 646 CCATGTTTAAATTCAGACACTCCAGTCAAGCAGTCAAGTGGAGTTGCTGTAATTT 705  
|||||  
DB 1 CCATGTTTAAATTCAGACACTCCAGTCAAGCAGTCAAGTGGAGTTGCTGTAATTT 60  
|||||  
DB 706 TGGCTCCATCTTTTTCACAGCAATGTAATTCGAAATTCGAAATTCGAAATTC 765  
|||||  
DB 61 TGGCTCCATCTTTTTCACAGCAATGTAATTCGAAATTCGAAATTCGAAATTC 120  
|||||  
DB 766 ATGGCTGAATTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 825  
|||||  
DB 121 ATGGCTGAATTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 180  
|||||  
DB 826 GTGAGAGTGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 885  
|||||  
DB 181 GTGAGAGTGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
|||||  
DB 886 AACTTCATTTTTCAGACAGTGTAAATGATGATGATGATGATGATGATGATGATGATG 945  
|||||  
DB 241 AACTTCATTTTTCAGACAGTGTAAATGATGATGATGATGATGATGATGATGATGATG 300  
|||||  
DB 946 ATCTGTTGCTTTCTGATACAGAGTGTGTTCC 980  
|||||  
DB 301 ATCTGTTGCTTTCTGATACAGAGTGTGTTCC 335  
|||||

## RESULT 15

ABV14061/c standard; cDNA: 406 BP.

XX ABV14061;

XX ABV14061;

DT 13-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker CDNA 14052.  
 XX  
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KM pharmacogenomic marker; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US05171.  
 XX  
 PR 17-FEB-2000; 2000US-183319P.  
 PR 16-MAR-2000; 2000US-189862P.  
 PR 25-MAY-2000; 2000US-207454P.  
 PR 09-JUN-2000; 2000US-211314P.  
 PR 18-JUL-2000; 2000US-219007P.  
 PR 13-DEC-2000; 2000US-255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 DR WPI; 2001-662795/76.  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer -  
 XX  
 PS Claim 1; Page 2343-2344; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 CC  
 XX  
 SQ Sequence 406 BP; 104 A; 99 C; 92 G; 111 T; 0 other;  
 XX  
 Query Match 18.8%; Score 264; DB 23; Length 406;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-28;  
 Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 518 CCTGGAATTCGTTTGGAAACCAAGTCTAGATGCAACGAATGTAGAACAGTCTT 577  
 DB 406 CCTGGAATTCGTTTGGAAACCAAGTCTAGATGCAACGAATGTAGAACAGTCTT 347  
 QY 578 CATGACGATGCGAGCTGAGATTAAAGGAAATGSGTCCCGAGACACAGCTGGTGTGC 637  
 DB 346 CATGACGATGCGAGCTGAGATTAAAGGAAATGSGTCCCGAGACACAGCTGGTGTGC 287  
 QY 638 TGAGAGTCCATGTTAAATTCAGAGCACTCCAGTCAAGCTGAGTGGAGTTCGTCG 697  
 DB 286 TGAGAGTCCATGTTAAATTCAGAGCACTCCAGTCAAGCTGAGTGGAGTTCGTCG 227  
 QY 698 CTAATATTTGCTCCATCTTTCTACAGCAATGAATTTGCAATCTGAACCAAGTGAA 757  
 DB 226 CTAATATTTGCTCCATCTTTCTACAGCAATGAATTTGCAATCTGAACCAAGTGAA 167  
 QY 758 AAAACAAATTCCTGAATGTAC 781

Db |||||  
 166 AAAACAAATTCCTGAATGTAC 143  
 Search completed: June 22, 2003, 19:53:56  
 Job time : 356 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 22, 2003, 19:33:58 ; Search time 1902 Seconds  
(without alignments) 11963.543 Million cell updates/sec

Title: US-09-820-003b-1

Perfect score: 1405  
Sequence: 1 aagcagatagctgagtcgcgc.....aaaaaaaaaaaaaaaaa 1405

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST.\*  
1: em\_estdb.\*  
2: em\_esthum.\*  
3: em\_estlin.\*  
4: em\_estlun.\*  
5: em\_estlun.\*  
6: em\_estlun.\*  
7: em\_estlun.\*  
8: em\_estlun.\*  
9: gb\_estl.\*  
10: gb\_estl.\*  
11: gb\_estl.\*  
12: gb\_estl.\*  
13: gb\_estl.\*  
14: gb\_estl.\*  
15: em\_estlun.\*  
16: em\_estlun.\*  
17: gb\_estl.\*  
18: em\_estlun.\*  
19: em\_estlun.\*  
20: em\_estlun.\*  
21: em\_estlun.\*  
22: em\_estlun.\*  
23: em\_estlun.\*  
24: em\_estlun.\*  
25: em\_estlun.\*  
26: em\_estlun.\*  
27: em\_estlun.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	743	52.9	948	9	AL539022
2	732.4	52.1	777	14	B0014597
3	723.8	51.8	811	14	AA746643
4	723.8	51.5	817	13	BI869965
5	708.8	50.4	910	14	B0276678
6	708.4	50.4	761	14	B0772048

7	707.8	50.4	945	9	AL530265
8	704.4	50.1	1105	13	BM452262
9	703.6	50.1	726	14	BM448090
10	702	50.1	766	13	B013092
11	692.8	49.3	760	14	BM71301
12	691.8	49.2	976	14	BM912800
13	690.6	49.2	837	13	BI086445
14	689	49.0	689	14	BM706159
15	688.2	49.0	737	13	BI092003
16	686	48.8	686	14	BM826571
17	685.6	48.8	717	9	AI952757
18	683.2	48.6	706	14	BM975245
19	682.6	48.2	796	13	BI659520
20	677.6	48.2	877	12	BG617277
21	676.4	48.1	678	13	BG939358
22	671.2	47.8	684	13	BI495589
23	670.8	47.7	1130	13	BM542833
24	669.6	47.7	813	14	B0179586
25	668.2	47.6	812	14	B0178972
26	661.2	47.1	686	14	BQ599300
27	659.6	46.9	721	9	AI884351
28	658	46.8	793	12	BG178268
29	657.6	46.8	690	10	AM150848
30	655.2	46.6	692	14	B0010358
31	650.4	46.3	692	9	AI828874
32	650	46.3	908	9	AI800097
33	648	46.1	723	9	AI858694
34	645.6	46.0	682	9	AI858694
35	640.4	45.6	900	12	BG032817
36	638.8	45.5	682	9	AI860577
37	638	45.4	638	14	BM783973
38	638	45.4	646	14	BO599347
39	638	45.4	687	10	BB646347
40	637.4	45.4	719	14	B0210709
41	632.8	45.0	1055	13	BM472577
42	631.8	45.0	664	10	BB617921
43	630.6	44.9	862	14	B0223273
44	630.6	44.9	645	10	AM166902
45	628.6	44.7	702	9	AI889183

#### ALIGNMENTS

RESULT 1  
LOCUS AL539022 948 bp mRNA linear EST 16-FEB-2001  
DEFINITION AL539022 LTI\_F1013.FBm1 Homo sapiens cDNA clone CS0DF030YG06 5  
ACCESSION AL539022  
VERSION AL539022.1 GI:12867866  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 948)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
source  
1..948  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DF030YG06"  
/clone\_1lb="LTI\_F1013.FBm1"  
/dev\_stage="pooled tissue from post conception fetuses (20 week - 24 week and 26 week)"  
/lab\_host="DH10B"



```

QY 613 GTCCCGAGACACACAGCTGTGTGTCTGAGAACGTAATGTAAATTCAGACACATCCAG 672
DB 718 GTCCCGAGACACACAGCTGTGTGTCTGAGAACGTAATGTAAATTCAGACACATCCAG 659
QY 673 TCAGACAGTCAGTGGAGAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 732
DB 658 TCAGACAGTCAGTGGAGAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 601
QY 733 AATTGCAATCTGAACCCCAAGTAAACCAAAATTCCTGAAATTCCTGATATGAGCT 792
DB 600 AATTGCAATCTGAACCCCAAGTAAACCAAAATTCCTGAAATTCCTGATATGAGCT 541
QY 793 GCACATCAACAGATTCCTACCGTCTCCACAAAGTCAGAGATGTAAATGTCAATACGT 852
DB 540 GCACATCAACAGATTCCTACCGTCTCCACAAAGTCAGAGATGTAAATGTCAATACGT 481
QY 853 ACCTTTTCTTTTATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 912
DB 480 ACCTTTTCTTTTATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 421
QY 913 CTTTGTGTCTGTATTAATAATATGTGTAAATCCTTTCTCTCTCTCTCTCTCTCTCTCT 972
DB 420 CTTTGTGTCTGTATTAATAATATGTGTAAATCCTTTCTCTCTCTCTCTCTCTCTCTCT 361
QY 973 TGTTCCTCGTGTGTGTAGATATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032
DB 360 TGTTCCTCGTGTGTGTAGATATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 301
QY 1033 GTCAAGTTAGTCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1092
DB 300 GTCAAGTTAGTCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 241
QY 1093 GTCTGTACATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152
DB 240 GTCTGTACATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181
QY 1153 GTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1212
DB 180 GTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 121
QY 1213 ACCTGTGATTAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1272
DB 120 ACCTGTGATTAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 61
QY 1273 CAATGCAATATGTTTAAATTAATTTGATCTTATTTGAAATGATGATGATGATGAT 1332
DB 60 CAATGCAATATGTTTAAATTAATTTGATCTTATTTGAAATGATGATGATGATGATGAT 1

RESULT 3
AA746643 811 bp mRNA linear EST 22-JAN-1998
LOCUS nx27c08.s1 NCI-CGAP.GC4 Homo sapiens cDNA clone IMAGE:1257326 3'
DEFINITION similar to gb:M28209 RAS-RELATED PROTEIN RAB-1A (HUMAN);, mRNA
sequence.
ACCESSION AA746643
VERSION AA746643.1 GI:2786629
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

```

DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLM at:  
[www.bio.linn.gov/birop/image/image.html](http://www.bio.linn.gov/birop/image/image.html)  
 Insert Length: 861 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 471.  
 Location/Qualifiers  
 1. 811  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:1257326"  
 /clone\_idb="NCI-CGAP\_GC4"  
 /tissue\_type="pooled germ cell tumors"  
 /lab\_host="DH10B"  
 /note="Vector: pUT3D-Pac (Pharmacia) with a modified  
 polylinker; 1st strand cDNA was prepared from 3 pooled  
 germ cell tumors, and was then primed with a Not I -  
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
 RI adaptors (Pharmacia), digested with Not I and cloned  
 into the Not I and Eco RI sites of the modified pUT3  
 vector. Library is normalized. Library was constructed by  
 Bento Soares and M. Fatima Bonaldo."

BASE COUNT 269 a 165 c 145 g 232 t  
 ORIGIN

Query Match. 51.8%; Score 727.6; DB 9; Length 811;  
 Best Local Similarity 95.8%; Pred. No. 9.8e-77;  
 Matches 780; Conservative 0; Mismatches 29; Indels 5; Gaps 3;

```

QY 502 AGCAATTTGCTGATTCCTCTGGAATTCCTTTTGGAAACGATGATGATGATGATGATGAT 561
DB 809 AGCGAATTTGCTGATTCCTCTGGAATTCCTTTTGGAAACGATGATGATGATGATGATGAT 752
QY 562 ATGTGAAGATGCTTTCATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 621
DB 751 ATGTGAAGATGCTTTCATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 695
QY 622 CAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681
DB 694 CAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 635
QY 682 CAGGTGAGAGTGTCTGTAATTTGCTGATTCCTTTTCTACAGCAATGATGATGATGAT 741
DB 634 CAGGTGAGAGTGTCTGTAATTTGCTGATTCCTTTTCTACAGCAATGATGATGATGATGAT 575
QY 742 TCTGAACCAAGTGAACCAAAATTCCTGATTTGATCTGATGATGATGATGATGATGATGAT 801
DB 574 TCTGAACCAAGTGAACCAAAATTCCTGATTTGATCTGATGATGATGATGATGATGATGAT 515
QY 802 CAGATTTCTACGCTCTCCACAAAGTCAGATGTAATGCTAATGCTAATGCTAATGCTAAT 861
DB 514 CAGATTTCTACGCTCTCCACAAAGTCAGATGTAATGCTAATGCTAATGCTAATGCTAAT 455
QY 862 TTAATTCCTTGACTCAAGACAGCTAATCTTATTTGATGATGATGATGATGATGATGATGAT 921
DB 454 TTAATTCCTTGACTCAAGACAGCTAATCTTATTTGATGATGATGATGATGATGATGATGAT 395
QY 922 CTGTGTTTAAATTAATGTTGTTATCTCTGTTGCTTCTGATGATGATGATGATGATGATGAT 981
DB 394 CTGTGTTTAAATTAATGTTGTTATCTCTGTTGCTTCTGATGATGATGATGATGATGATGAT 335
QY 982 TGTGTGTGTAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1041
DB 334 TGTGTGTGTAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 275
QY 1042 ACTGTCTGAAGATGAAGTCAAGCATTTTGTATCAACAGACAGCATGCTGTCTGAC 1101
DB 274 ACTGTCTGAAGATGAAGTCAAGCATTTTGTATCAACAGACAGCATGCTGTCTGAC 215
QY 1102 TTTCCATGATTAATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
DB 214 TTTCCATGATTAATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 155

```

QY 1162 TGTAGAGTTAATGAAAGATTACACTATCTGATTATAGTTCTTCATCTCTGCAT 1221  
 DB 154 TGTAGAGTTAATGAAAGATTACACTATCTGATTATAGTTCTTCATCTCTGCAT 95  
 QY 1222 ATATTTTGGCTGACAAATTTGTAATTTGTGACACTATGTACAAAACACTGAG 1281  
 DB 94 ATATTTTGGCTGACAAATTTGTAATTTGTGACACTATGTACAAAACACTGAG 35  
 QY 1282 ATATTTTAAATAATATGTTACTTATTTGGAAGTA 1315  
 DB 34 ATATTTTAAATAATATGTTACTTATTTGGAAGTA 1

RESULT 4  
 LOCUS B1869965 817 bp mRNA linear EST 11-OCT-2001  
 DEFINITION 603394116P1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:5404275 5',  
 mRNA sequence.  
 ACCESSION B1869965 GI:16043638  
 VERSION B1869965.1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.lnl.gov  
 Plate: L14M12030 row: P column: 04  
 High quality sequence stop: 669.

FEATURES  
 source Location/Qualifiers  
 1..817

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="5404275"  
 /clone\_lib="NIH\_MGC\_90"  
 /issue\_type="adenocarcinoma, cell line"  
 /note="Organ: liver; Vector: pCMVSPORT6; Site: 1: NotI;  
 Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC library."

BASE COUNT 237 a 155 c 169 g 256 t.

Query Match 51.5%; Score 723.8; DB 13; Length 817;  
 Best Local Similarity 95.3%; Pred. No. 2.7e-76;  
 Matches 779; Conservative 0; Mismatches 32; Indels 6; Gaps 3;

QY 475 AGAAGTGTAGTACACACACAGGAAGATTGCTGATTCCTTGAATTCGCTTT 534  
 DB 1 AGAAGTGTAGTACACACACAGGAAGATTGCTGATTCCTTGAATTCGCTTT 60  
 QY 535 TGAAGACAGTGTAGTACACACAGGAAGATTGCTGATTCCTTGAATTCGCTTT 594  
 DB 61 TGAAGACAGTGTAGTACACACAGGAAGATTGCTGATTCCTTGAATTCGCTTT 120  
 QY 595 AGATTAAAGAGGATGGGTCGAGACAGGACAGTGTGTGTGAGAGTCCATGTTA 654  
 DB 121 AGATTAAAGAGGATGGGTCGAGACAGGACAGTGTGTGTGAGAGTCCATGTTA 180

QY 655 AAATTCAGACAGCTCCAGCTCAAGAGTCAAGGTTGCTGCTAAATTTGCCCTCAT 714  
 DB 181 AAATTCAGACAGCTCCAGCTCAAGAGTCAAGGTTGCTGCTAAATTTGCCCTCAT 240  
 QY 715 CCTTTTCACAGCAATTCATTTTCAATCTCAACCAAGTCAAAAAACAAATTTGCCCTGA 774  
 DB 241 CCTTTTCACAGCAATTCATTTTCAATCTCAACCAAGTCAAAAAACAAATTTGCCCTGA 300  
 QY 775 ATTGACAGTGTAGTACGACACTCAACAGATTCCTACCGCTCCACAAAGTCAAGAT 834  
 DB 301 ATTGACAGTGTAGTACGACACTCAACAGATTCCTACCGCTCCACAAAGTCAAGAT 360  
 QY 835 TGTAAATGTCATACACTGCTTTTATTTTATTCCTGACTCAAGACAGTCACTCAT 894  
 DB 361 TGTAAATGTCATACACTGCTTTTATTTTATTCCTGACTCAAGACAGTCACTCAT 420  
 QY 895 TTGCAAGTGTTTAAACCTTTGTGTGTGTGTGTATATAAATGTGTATCTCTGT 954  
 DB 421 TTGCAAGTGTTTAAACCTTTGTGTGTGTGTGTATATAAATGTGTATCTCTGT 480  
 QY 955 GCTTTCCGATACAGAGCTGTTCCCGGTGGTGTAGATAATATTTTGTGTATGTT 1014  
 DB 481 GCTTTCCGATACAGAGCTGTTCCCGGTGGTGTAGATAATATTTTGTGTATGTT 540  
 QY 1015 ATATTGCGATTTAGATGTGAGTTTATGCTTCTGAAAGTGAAGTCACTTTGTA 1074  
 DB 541 ATATTGCGATTTAGATGTGAGTTTATGCTTCTGAAAGTGAAGTCACTTTGTA 600  
 QY 1075 TCAACAGCAGCAGAGGTCGTCACCTTCATGACATTAAGTTAGAGATGTTAT 1134  
 DB 601 TCAACAGCAGCAGAGGTCGTCACCTTCATGACATTAAGTTAGAGATGTTAT 660  
 QY 1135 GTAAAGTCT--GATTGCTAGTCTTCTTCTGTGTAGATTATTAATGGAAGATTACATA 1191  
 DB 661 GTAAAGTCTGATTGCTAGTCTTCTTCTGTGTAGATTATTAATGGAAGATTACATA 720  
 QY 1192 TCTGATTATATCT-TTCTCACTCTGCAATATTAATTTGGCTGCAGAA--TATTGTAA 1248  
 DB 721 TCTGATTATATCTTCTCTCACTCTGCAATATTAATTTGGCTGCAGAA--TATTGTAA 780  
 QY 1249 TTTGTTGACACTATGTAAACAAACACTGAAATAT 1285  
 DB 781 TATGTGCCACTATGTAAACAAACACTGAAATAT 817

RESULT 5  
 LOCUS BQ276678 910 bp mRNA linear EST 07-MAY-2002  
 DEFINITION AGENCOURT\_7025320 NIH\_MGC\_126 Homo sapiens cDNA clone IMAGE:5809629  
 5', mRNA sequence.  
 ACCESSION BQ276678  
 VERSION BQ276678.1 GI:20486886  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: NCI  
 cDNA Library Preparation: Michael Brownstein Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.lnl.gov  
 Plate: LICM205 row: 1 column: 22  
 High quality sequence stop: 558.

FEATURES  
 Location/Qualifiers

```

source
1. .910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="NIH_MGC_126"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (TI-phage-resistant)"
/note="Vector: pDRN-LIB; Site_1: Sfil (ggccattatgccc);
Site_2: Sfil (ggccgcctggccc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon
- 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-ATGCGATGATCAACGACGAGTGGCCATTCAGCGCGG-3' and
5'-ATTCGAGAGCCGCGGCGCGCATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5-1
kb size fraction (other fractions present in NIH_MGC_127
and NIH_MGC_128). Library created in the laboratory of T.
Usdin, M.D., Ph.D. (NIH). Note: this is a NIH_MGC
Library."
BASE COUNT      243 a      194 c      180 g      292 t      1 others
ORIGIN
Query Match      50.4%; Score 708.8; DB 14; Length 910;
Best Local Similarity 98.9%; Pred. No. 1.4e-74;
Matches 724; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

OY 584 GATGCGACGATTAATAAAGCAATGGTCCCGACCAACGCGGCGCGAGAA 643
    |||||||
DB 3 GATGCGACGATTAATAAAGCAATGGTCCCGACCAACGCGGCGCGAGAA 62
    |||||||
OY 644 GTCGCAATTAATAATTCAGACACTCCAGTCAAGACAGTGGAGTGGCTCTTAAA 703
    |||||||
DB 63 GTCGCAATTAATAATTCAGACACTCCAGTCAAGACAGTGGAGTGGCTCTTAAA 122
    |||||||
OY 704 TTTCGCTCCATCTTTCTCAGACAGTAATTTGCAATCCCAAGTGAAGAAAAA 763
    |||||||
DB 123 TTTCGCTCCATCTTTCTCAGACAGTAATTTGCAATCCCAAGTGAAGAAAAA 182
    |||||||
OY 764 AAATTCGCTTAATGCTAGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823
    |||||||
DB 183 AAATTCGCTTAATGCTAGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
    |||||||
OY 824 AGGTTCAGAGATTTGAATGCTAATGCTAATGCTAATGCTAATGCTAATGCTAATGCT 883
    |||||||
DB 243 AGGTTCAGAGATTTGAATGCTAATGCTAATGCTAATGCTAATGCTAATGCTAATGCT 302
    |||||||
OY 884 CTAACTTCATTTTCAAGACTGTTTAAACCTTGTGTCTGCTGCTGCTGCTGCTGCTGCT 943
    |||||||
DB 303 CTAACTTCATTTTCAAGACTGTTTAAACCTTGTGTCTGCTGCTGCTGCTGCTGCTGCT 362
    |||||||
OY 944 TTAATCTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
    |||||||
DB 363 TTAATCTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
    |||||||
OY 1004 TTTTGAATGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
    |||||||
DB 423 TTTTGAATGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
    |||||||
OY 1064 GGCATTTTGTATCAACAGACAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123
    |||||||
DB 483 GGCATTTTGTATCAACAGACAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
    |||||||
OY 1124 AGATGTTAATGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1183
    |||||||
DB 543 AGATGTTAATGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
    |||||||
OY 1184 TTACACTCTGATTAATGTTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1243
    |||||||

```

```

DB 603 TTACACTCTGATTAATGTTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662
OY 1244 TGTAAATTTTGGCAGCTACTGTATACAAAACAGTGAAGATA-TGTTAATAAATATGTA 1302
    |||||||
DB 663 TGTAAATTTTGGCAGCTACTGTATACAAAACAGTGAAGATA-TGTTAATAAATATGTA 722
    |||||||
OY 1303 CTTAATGGAAGT 1314
    |||||||
DB 723 CTTAATGGAAGT 734
    |||||||

RESULT 6
BO772048/C
LOCUS
DEFINITION
BO772048 761 bp mRNA linear EST 26-JUL-2002
UT-H-EZ1-bdk-1-10-0-UI.s1 NCI-CGAP-Ch2 Homo sapiens cDNA clone
UT-H-EZ1-bdk-1-10-0-UI 3', mRNA sequence.
BO772048
BO772048.1 GI:21980524
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 761)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cygpb@remail.nih.gov
Tissue Procurement: Dr. Steven Citelis/ Rush Presbyterian, Dept. of
Orthopaedics
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
CDNA Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento.soares@uiowa.edu
Seq primer: M13 FORWARD
POLY-A-Tes.

FEATURES
Location/Qualifiers
1..761
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UT-H-EZ1-bdk-1-10-0-UI"
/clone_1lb="NCI-CGAP-Ch2"
/tissue_type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Left Pelvis; Vector: p773-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI-CGAP-Ch2 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma Grade II. The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into p773-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TGATCAGCGT.
TAG_LIB-UI-H-EZ1
TAG_TISSUE-gr-2-chondrosarcoma
TAG_SEQ-ATCTAATGAG"
BASE COUNT      258 a      143 c      125 g      235 t
ORIGIN
Query Match      50.4%; Score 708.4; DB 14; Length 761;
Best Local Similarity 99.3%; Pred. No. 1.0e-74;
Matches 753; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

OY 575 TTTCATGACGATGCGACGCTGAGATTAAAGCAATGGTCCCGACCAACAGCTGGTGG 634

```



Db	754	TTTCATGACGAT - GCAGCTGAGATTAAAAACGC - ATTGTCCGGAGCAACAGCTGCT - G	698
QY	635	TGCTGACAAAGTCACATGTTAAATTCAGAGACATCCAGCTCAAGCAGTAGGTGAGGTG	694
Db	697	TGCTGAGAGTCACATGTTAAATTCAGAGACATCCAGCTCAAGCAGTAGGTGAGGTG	638
QY	695	CTGCTAAATTTGGCTCCATCCCTTTTCACAGACAATGAATTTGGAACTCAACCCAGT	754
Db	637	CTGCTAAATTTGGCTCCATCCCTTTTCACAGACAATGA - TTGCAATCTGAACCCAGT	579
QY	755	GAATAAAACAAATTTGCTGAACTGTGACTGTATGTAGCTGCACCTACACAGATTTCTTACG	814
Db	578	GAATAAAACAAATTTGCTGAACTGTGACTGTATGTAGCTGCACCTACACAGATTTCTTACG	519
QY	815	TCCTCACAAAGGTGAGAGATGTAAATGTCGAATACACTTTTTTTTATTCCTGTAC	874
Db	518	TCCTCACAAAGGTGAGAGATGTAAATGTCGAATACACTTTTTTTTATTCCTGTAC	459
QY	875	TCAGACAGCTACACTATTTTCAGAACTGTTTAAACCTTTGTCGTGTTTATAAA	934
Db	458	TCAGACAGCTACACTATTTTCAGAACTGTTTAAACCTTTGTCGTGTTTATAAA	399
QY	935	TAATGTGTATCTCTGTGCTTTTCTGTATACAGACTGTTTCCGTGTTGTTAGAA	994
Db	398	TAAAGTGTATCTCTGTGCTTTTCTGTATACAGACTGTTTCCGTGTTGTTAGAA	339
QY	995	TATATTTGTTTATGTTTATTTATTTGGCAGTTTATAGATGTGAGGTTTGTCTTGAGA	1054
Db	338	TATATTTGTTTATGTTTATTTATTTGGCAGTTTATAGATGTGAGGTTTGTCTTGAGA	279
QY	1055	TGAAGTTCAGCAATTTGTATCAAAACAGACACAGCAAGTGTGTGCATTTCCATGCATTA	1114
Db	278	TGAAGTTCAGCAATTTGTATCAAAACAGACACAGCAAGTGTGTGCATTTCCATGCATTA	219
QY	1115	AGTTATGAGATGTTTATATGTATAGATGTGATTTGCTATCTTCTGTAGAGTTATA	1174
Db	218	AGTTATGAGATGTTTATATGTATAGATGTGATTTGCTATCTTCTGTAGAGTTATA	159
QY	1175	ATGGAAGATTTACACTATCTATCTATTAATAGTTTCTTCACACTGTGCATATTAATTTGGCT	1234
Db	158	ATGGAAGATTTACACTATCTATCTATTAATAGTTTCTTCACACTGTGCATATTAATTTGGCT	99
QY	1235	GCACAATATTTGTAATTTGTTCACACATGTGAACAAAACAACCTGAAGATATGTTAATA	1294
Db	98	GCACAATATTTGTAATTTGTTCACACATGTGAACAAAACAACCTGAAGATATGTTAATA	39
QY	1295	ATATTTGACTATTTGGAAGTAAAAAATTTTTTTTTT 1332	
Db	38	ATATTTGACTATTTGGAAGTAAAAAATTTTTTTTTT 1	

RESULT 7	AL530265	945 bp	mRNA	linear	EST 13-FEB-2001
LOCUS	AL530265				
DEFINITION	AL530265 LTL.NEUF001.NBC4 Homo sapiens cDNA clone CSDD009YH19 5				
ACCESSION	AL530265				
VERSION	AL530265.1	GI:12793758			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1. (bases 1 to 945)				
AUTHORS	Ll.M.B., Gruber,C., Jessee,J. and Polayes,D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: <a href="mailto:secret@genoscope.cns.fr">secret@genoscope.cns.fr</a> , Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> .				
FEATURES	Location/Qualifiers				

```

source
1. .945
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD0091H19"
/clone_11b="LT1_NFL001_NBC4"
/sex="male"
/issue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by life technologies. Contact : Feng liang life
technologies, a division of invitrogen 9800 Medical Centre
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com"
http://fulllength.invitrogen.com"
BASE COUNT      280 a      186 c      224 g      246 t      9 others
ORIGIN

Query Match      50.4%; Score 707.8; DB: 9; Length 945;
Best Local Similarity 88.4%; Pred. No. 1.8e-74;
Matches 835; Conservative 5; Mismatches 7; Indels 98; Gaps 3

QY      60 CGTTTCCTTCCCGGAGACAGCCTATCTATCTCTTCTTGATGATACCTCGCGGGAG 119
      1 CGTTTCCTTCCCGGAGAGCCTATCTATCTCTTCTTGATGATACCTCGCGGGAG 60

QY      120 AGTCAGGCGGCGGCGTGGCGGAGCAAGGCGCGGTGGCGGCGGCGGACCTCATGAC 179
      61 AGTCAGGCGGCGGCGTGGCGGAGCAAGGCGCGGTGGCGGCGGCGGACGACATGAC 120

QY      180 ATGTCCAGATAAATCCGAAATATATATTTATTTCAAGTACTCTTGATGGCGACTA 239
      121 ATGTCCAGATAAATCCGAAATATATATTTATTTCAAGTACTCTTGATGGCGACTA 180

QY      240 GGGGTGGAAAGTCTTCCTTCTTGTAGGTTTGAGTGAGTATACAGAAAGCTAC 299
      181 GGGGTGGAAAGTCTTCCTTCTTGTAGGTTTGAGTGAGTATACAGAAAGCTAC 240

QY      300 ATCAGACAAATGGTGATTTCAAATTAAGACTATAGGTAGAGGGGAAACATC 359
      241 ATCAGACAAATGGTGATTTCAAATTAAGACTATAGGTAGAGGGGAAACATC 300

QY      360 AAGCTTCAAAT----- 371
      301 AAGCTTCAAATATGGGAGACACAGACAGGCAAGAAAGATTGGAACAATCAGCTCAAGTTAT 360

QY      372 -----GAGTCCCTCAAT 383
      361 TACAGAGAGCCCATGGCATCATAGTTGTATGATGTGACAGATCAGAGTCCCTCAAT 372

QY      384 AATGTTAAACAGTGGCTGCAGGAATATGATGCTATGCGAGTGAATATGCAACAATG 443
      421 AATGTTAAACAGTGGCTGCAGGAATATGATGCTATGCGAGTGAATATGCAACAATG 480

QY      444 TTGGTAGGAGACAAATGTGATCTGACCAAAAGAAAGTAGACTACACAGCGAAG 503
      481 TTGGTAGGAGACAAATGTGATCTGACCAAAAGAAAGTAGACTACACAGCGAAG 540

QY      504 GAATTTGCTGATTCCTTGGAAATTCGTTTTTGGAAACAGTGTAAAGATGCAAGAAAT 563
      541 GAATTTGCTGATTCCTTGGAAATTCGTTTTTGGAAACAGTGTGTAAAGATGCAAGAAAT 600

QY      564 GTAGAACACTCTTTCATGACGATGCGACCTAGATTTAAAAAGCAATGGGTCCGGAGCA 623
      601 GTAGAACACTCTTTCATGACGATGCGACCTAGATTTAAAAAGCAATGGGT-CCGGAGCA 659

QY      624 ACAAGTGTGTGCTGAGAAAGTCCAAATTTAAATTCAGAGCACTCCAGTCAAGCACTGA 683
      660 ACAAGTGTGTGCTGAGAAAGTCCAAATTTAAATTCAGAGCACTCCAGTCAAGCACTGA 719

```

QY 684 -GGTGGAGCTTGCCTGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 742  
DB 720 GGGGTGAGCTTGCCTGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 779  
QY 743 CTGAACCCAGTATGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 802  
DB 780 CTGAACCCAGTATGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 839  
QY 803 AGATTCTTACCGTCTCCAGCAAGTATGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAG 862  
DB 840 AGATTCTTACCGTCTCCAGCAAGTATGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAG 899  
QY 863 TATTCCTTGCCTGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 907  
DB 900 TATTCCTTGCCTGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 944

RESULT 8  
BM452262 1105 bp mRNA linear EST 05-FEB-2002  
LOCUS BM452262  
DEFINITION AGENCOURT\_6386191 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:552635  
5', mRNA sequence.  
ACCESSION BM452262.1 GI:18501302  
VERSION BM452262.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1105)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph. D.  
Email: cgsrbs-remail.nih.gov  
Tissue Procurement: ATCC/DOCTD/DP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LAM12200 row: 3 column: 12  
High quality sequence stop: 733.  
Location/Qualifiers  
1..1105  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="552635"  
/clone\_id="NIH\_MGC\_72"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pCMV-SPORT6; site\_1: NotI;  
site\_2: SalI; Cloned unidirectionally. Primer: Oligo dr.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

BASE COUNT 320 a 216 c 279 g 282 t 8 others  
ORIGIN

Query Match 50.1%; Score 704.4; DB 13; Length 1105;  
Best Local Similarity 86.0%; Pred. No. 4.1e-74;  
Matches 899; Conservative 0; Mismatches 43; Indels 103; Gaps 7;

QY 1 AAGCATAGCTGAGT-GCGGGGCTGCTGATTTGTTCTAGGGGAGGAGTAGGGGAGA 59  
DB 2 AAGCATAGCTGAGTGGGGGGCTGCTGATTTGTTCTAGGGGAGGAGTAGGGGAGA 61  
QY 60 GCTTGGCTTCCGGAAGACGCTATGCTATTTCTTTGATTAACCGTGGCGGAG 119  
DB 62 GCTTGGCTTCCGGAAGACGCTATGCTATTTCTTTGATTAACCGTGGCGGAG 121  
QY 120 AGTCAGGGCGGCGGCGGAGCAAGGGCGGCGGCGGCGGCGGCGGCGGAG 179  
DB 120 AGTCAGGGCGGCGGCGGAGCAAGGGCGGCGGCGGCGGCGGCGGCGGAG 179

DB 122 AGTCAGGGCGGCGGCGGAGCAAGGGCGGCGGCGGCGGCGGCGGCGGAGCTGACATGC 181  
QY 180 ATGTCAGGATTAATCCCGAATATGATTTATTTATTCAGATTTCTGATGGGACTCA 239  
DB 182 ATGTCAGGATTAATCCCGAATATGATTTATTTATTCAGATTTCTGATGGGACTCA 241  
QY 240 GGGGTGGAAGCTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 299  
DB 242 GGGGTGGAAGCTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 301  
QY 300 ATCAGCAATTTGTTGATTTTCAATTAAGAACTATAGTTAGAGGGGAAAGCAATC 359  
DB 302 ATCAGCAATTTGTTGATTTTCAATTAAGAACTATAGTTAGAGGGGAAAGCAATC 361  
QY 360 AAGCTTCAATTA----- 371  
DB 362 AAGCTTCAATTAATGAGACAGACGCGGAGGAAAGATTTGGAACATCCTCAGTAT 421  
QY 372 -----GATCCTTCAT 383  
DB 422 TACAGAGAGCCCATGCGATCATATGTTGATGATGTGACAGATCAGAGATCCTTCAT 481  
QY 384 AATGTAACAGTGGCTGAGCAAGTAATGATGATGATGATGATGATGATGATGATGATGATG 443  
DB 482 AATGTAACAGTGGCTGAGCAAGTAATGATGATGATGATGATGATGATGATGATGATGATG 541  
QY 444 TTGGTAGGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 503  
DB 542 TTGGTAGGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601  
QY 504 GAATTTGCTGATTCCTTGGAAATTCCTTTTGGAAACAGGCTGTAAGAAAGCAAGAT 563  
DB 602 GAATTTGCTGATTCCTTGGAAATTCCTTTTGGAAACAGGCTGTAAGAAAGCAAGAT 661  
QY 564 GTGAACACTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 623  
DB 662 GTGAACACTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721  
QY 624 ACAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 683  
DB 722 ACAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780  
QY 684 -GGGAGGCTTCTGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 742  
DB 781 NGGTGAGGCTTCTGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 840  
QY 743 CTGAACCCAGTATGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 802  
DB 841 TGAACCCAGTATGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAGTAATTTGCAAT 899  
QY 803 AGATTCTTACCGTCTCCAGCAAGTATGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAG 860  
DB 900 AGATTCTTACCGTCTCCAGCAAGTATGTAATAATTTGCGTCCATCCCTTTCTCAGAGCAAG 959  
QY 861 TTTATTTCCCTTACCAAGACACTTATTTTCAAGAGCTTTTAAACCTTTTGG 919  
DB 960 TTTATTTCCCTTACCAAGACACTTATTTTCAAGAGCTTTTAAACCTTTTGG 1019  
QY 920 TCGTGGTTATTAATATGTTGT 944  
DB 1020 TCGTGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1044

RESULT 9  
B0448090/c 725 bp mRNA linear EST 28-MAY-2002  
LOCUS B0448090  
DEFINITION UT-H-FU1-bai-k-14-0-UT-51 NCI-GAP-CT1 Homo sapiens cDNA clone  
UT-H-FU1-bai-k-14-0-UT-3', mRNA sequence.  
ACCESSION B0448090  
VERSION B0448090.1 GI:21251202  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 725)  
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: MJ3 FORWARD  
POLYA-Yes.

## FEATURES

Source  
Location/Qualifiers

1..725  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-H-E01-Dal-K-14-0-UI"  
/clone.lib="NCI-CCGAP.CLI"  
/tissue.type="Osteoarthritic Cartilage"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Knee; Vector: pTV73-Pac (Pharmacia) with a modified polylinker; Site.1: Ecor I; Site.2: Not I; NCI-CCGAP.CLI is a normalized cDNA library containing the following tissue(s): Osteoarthritic Cartilage The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pTV73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TAGTCACGCT.  
TAG\_LIB=UI-H-E01  
TAG\_TISSUE="Osteoarthritic cartilage"  
TAG\_SEQ="TAGTCACGCT"  
BASE COUNT 247 a 135 c 120 g 221 t 2 others  
ORIGIN

Query Match 50.1%; Score 703.6; DB 14; Length 725;  
Best Local Similarity 99.0%; Pred. No. 6.8e-74;  
Matches 717; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

609 ATGGGTCCCGAGCAACAGCTGGTGGTCTGAGAAAGTCAATGTTAAATTCAGAGCACT 668  
|||||  
723 AGGGGTCCCGAGCAACAGCTGTGTTNGCTGAGAGTCCAAATGTTAAATTCAGAGCACT 664  
|||||  
669 CCAGTCACAGCTGAGAGTGGTGGTCTGAGAAATTTGCCATCCCTTTCTCAGACG 728  
|||||  
663 CCAGTCACAGCTGAGAGTGGTGGTCTGAGAAATTTGCCATCCCTTTCTCAGACG 605  
|||||  
729 AATGATTTGCAATCTGAACCAAGTGAACCAAAATTTGCTTAATTTGCTAGTGT 788  
|||||  
604 AATGATTTGCAATCTGAACCAAGTGAACCAAAATTTGCTTAATTTGCTAGTGT 545  
|||||  
789 AGCTCAGCTACCAACAGATTTTACCGCTCTCCACAAAGGTGAGAGATTGTAATGGTCAAT 848  
|||||  
544 AGCTCAGCTACCAACAGATTTTACCGCTCTCCACAAAGGTGAGAGATTGTAATGGTCAAT 485  
|||||  
849 ACTGACTTTTATATCCCTTGACCTCAAGACAGCTAATCTTATTCAGAACTGTTT 908  
|||||  
484 ACTGACTTTTATATCCCTTGACCTCAAGACAGCTAATCTTATTCAGAACTGTTT 425  
|||||  
909 AAACCTTTGTGCTGTTTATTAATATGTTGTTAATCCCTTTGCTTCTGATATCC 968  
|||||  
424 AAACCTTTGTGCTGTTTATTAATATGTTGTTAATCCCTTTGCTTCTGATATCC 365

QY 969 AGACTGTTTCCCGTGGTGGTGGTGAATATATTTGTTGATGTTAATTTGGCAATGTTT 1028  
|||||  
Db 364 AGACTGTTTCCCGTGGTGGTGGTGAATATATTTTGTGATGTTAATTTGGCAATGTTT 305  
|||||  
QY 1029 AGATGTCAAGTTTACTCTTCTGACAGTGAAGTTCAGCCATTTTGTATCAACAGCAAG 1088  
|||||  
Db 304 AGATGTCAAGTTTACTCTTCTGACAGTGAAGTTCAGCCATTTTGTATCAACAGCAAG 245  
|||||  
QY 1089 CAGTGTCTGCTACTTCCATGATGATTAAGTTAGTGAATGTAATGTAATGTAATGATTT 1148  
|||||  
Db 244 CAGTGTCTGCTACTTCCATGATGATTAAGTTAGTGAATGTAATGTAATGTAATGATTT 185  
|||||  
QY 1149 GCTACTTCTCTCTTGTAGAGTTAATATGAAAGATTACATCTATCTATTAATGTTCT 1208  
|||||  
Db 184 GCTACTTCTCTCTTGTAGAGTTAATATGAAAGATTACATCTATCTATTAATGTTCT 125  
|||||  
QY 1209 TCATCTCTGCTATTAATTTTGTGCTGCAAGATTAATTTGTTGCACTATGTAAC 1268  
|||||  
Db 124 TCATCTCTGCTATTAATTTTGTGCTGCAAGATTAATTTGTTGCACTATGTAAC 65  
|||||  
QY 1269 AAAACACTGAAGATGTTTATTAATATTTGATTTTGGAGTAAAAAAGAAAAA 1328  
|||||  
Db 64 AAAACACTGAAGATGTTTATTAATATTTGATTTTGGAGTAAAAAAGAAAAA 5  
|||||  
QY 1329 AAAA 1332  
|||||  
Db 4 AAAA 1

RESULT 10  
BI913092  
LOCUS  
DEFINITION  
603179695F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5244094.5',  
mRNA sequence.  
BI913092  
ACCESSION  
BI913092.1 GI:16177393  
VERSION  
BI913092.1  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 796)  
NIH-MGC <http://mhc.ncbi.nlm.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM11615 row: m column: 23  
High quality sequence stop: 784.

## FEATURES

Location/Qualifiers

1..796  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5244094"  
/clone.lib="NIH\_MGC\_121"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcorV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcorV site is destroyed upon cloning). Average insert size 1.7 kb. Insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note:

BASE COUNT 235 a 140 c 168 g 253 t  
 ORIGIN

this is a NIH-MGC library \*

Query Match 50.0%; Score 702; DB 13; Length 796;  
 Best Local Similarity 95.7%; Pred. No. 9, 8e-74;  
 Matches 734; Conservative 0; Mismatches 30; Indels 4; Gaps 3;

485 AGACTACACACAGGAGAAATTTGCTGATTCCTTGGAAATTCCTTTTGGAAACCG 544  
 1 AGACTACACACAGGAGAAATTTGCTGATTCCTTGGAAATTCCTTTTGGAAACCG 57  
 545 TGTCTAGATGACAGCAATGTAGAACAGTCTTTCATGACAGTGCAGTGCAGATTTAAAA 604  
 58 TGTCTAGATGACAGCAATGTAGAACAGTCTTTCATGACAGTGCAGTGCAGATTTAAAA 117  
 605 GCGAATGGTCCCGAGACAGCAGTGGTGGTCTGAGAAAGTCCAAATGTTAAAAATTCAG 664  
 118 GCGAATGGTCCCGAGACAGCAGTGGTGGTCTGAGAAAGTCCAAATGTTAAAAATTCAG 177  
 665 CACTTCAGTCACAGCAGTGCAGTGGTGGTCTGAGAAATTTTCCATCCCTTTCTCA 724  
 178 CACTTCAGTCACAGCAGTGCAGTGGTGGTCTGAGAAATTTTCCATCCCTTTCTCA 237  
 725 CAGCAATGATTTGCAATCTGAACCCAGTGAAGAA-AAAAGAAATTCCTGATTTGCTG 783  
 238 CAGCAATGATTTGCAATCTGAACCCAGTGAAGAAAGAAATTCCTGATTTGCTG 297  
 784 TATGTAGCTGACATCAACAGATTTTACCGCTCCACAAAGGTCCAGATTTGTAATG 843  
 298 TATGTAGCTGACATCAACAGATTTTACCGCTCCACAAAGGTCCAGATTTGTAATG 357  
 844 TCATATGACGCTTTTATTTTATTCCTGACACAGCAGTGCATCTTTCATTTGCAAGCT 903  
 358 TCATATGACGCTTTTATTTTATTCCTGACACAGCAGTGCATCTTTCATTTGCAAGCT 417  
 904 GTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963  
 418 GTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 477  
 964 ATACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023  
 478 ATACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537  
 1024 TATCTAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1083  
 538 TATCTAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597  
 1084 ACAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1143  
 598 ACAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657  
 1144 GATTTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1203  
 658 GATTTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717  
 1204 TTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1263  
 718 TTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777  
 1264 GTACAAA 1271  
 778 GTACAAA 785

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 760)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtriction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: McCray, PB  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
 Seq primer: M13 FORWARD  
 POLYA=yes.  
 FEATURES  
 source location/Qualifiers  
 1..760  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UT-CF-BC1-ab1-a-17-0-UT"  
 /clone\_1lb="UT-CF-BC1"  
 /tissue\_type="lung"  
 /dev\_stage="Adult and fetal"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /note="Organ: lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site\_1: Ecor I; Site\_2: Not I; UT-CF-BC1 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)18 tail. The sequence tag for this library is AAGTCCTTAC.  
 TAG\_LIB-UT-CF-BC1  
 TAG\_TISSUE=Normal lung Epithelial Cells Tissue nos 369-371 and 380-383  
 TAG\_SEQ=AAGTCCTTAC"  
 BASE COUNT 255 a 142 c 125 g 236 t 2 others  
 ORIGIN  
 Query Match 49.3%; Score 692.8; DB 14; Length 760;  
 Best Local Similarity 99.1%; Pred. No. 1.2e-72;  
 Matches 738; Conservative 0; Mismatches 3; Indels 4; Gaps 4;  
 588 GCGAGGAGATTAAAGGAGTGGGTCGCGAGCAACAGCTGGTGTGAGAAAGCC 647  
 741 GCGAGGAGATTAAAGGAGTGGGTCGCGAGCAACAGCT-GTGTGAGTGAAGATCC 684  
 648 AATGTTAAATTTGAGAGCACTCCAGTCAAGTGCAGTGGTGGTGGTGGTGGTGGTGG 707  
 683 AATGTTAAATTTGAGAGCACTCCAGTCAAGTGCAGTGGTGGTGGTGGTGGTGGTGG 766  
 708 CCGTCATCTTTTTCACAGCAAGTAATTTGCAATGTGACCAAGTGAAGAAACAAAT 767  
 625 CCGTCATCTTTTTCACAGCAAGTAATTTGCAATGTGACCAAGTGAAGAAACAAAT 566  
 768 TGCTGAAATTTGCTGATGTGAGTGCACACAGATTTTACGCTCTCCCAAGCT 827

```

|||||
565 TCCGGAATGCTACGTATAGTGCACATACAGAGATCTACCGCTCCACAAAGCT 506
828 CAGAGATGTAATGGTCAATGCTTTTATCCCTGACCTCAAGACCTAA 887
505 CAGAGATGTAATGGTCAATGCTTTTATCCCTGACCTCAAGACCTAA 446
888 CTTCAATTTTCAAGACTGTTTAACTTTGCTGCTGTTTAAATATGCTGTAAT 947
445 CTTCAATTTTCAAGACTGTTTAACTTTGCTGCTGTTTAAATATGCTGTAAT 386
948 CTTGCTGCTTCCGATACGACGCTTTCCCGTGGTTGTTAGATATATTTGTTT 1007
385 CTTGCTGCTTCCGATACGACGCTTTCCCGTGGTTGTTAGATATATTTGTTT 326
1008 GATGTTATATGGCATGTTTATGATGCTGATGCTTCAAGATATAGTCCGCA 1067
325 GATGTTATATGGCATGTTTATGATGCTGATGCTTCAAGATATAGTCCGCA 266
1068 TTTTGTATCAACAGACAGACAGTGTCTGCTCACTTCATGCAATTAAGTTAGTAGAT 1127
265 TTTTGTATCAACAGACAGACAGTGTCTGCTCACTTCATGCAATTAAGTTAGTAGAT 206
1128 GTTATATGTAATGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187
205 GTTATATGTAATGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146
1188 ACTATCTGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1247
145 ACTATCTGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 86
1248 ATTGTTGTCACACTATGTAACAAACAACTGGAAGATATGTTAAATATGCTGCTGCT 1307
85 ATTGTTGTCACACTATGTAACAAACAACTGGAAGATATGTTAAATATGCTGCTGCT 26
1308 TGGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1332
25 TGGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1

```

RESULT 12  
BM919860  
LOCUS  
DEFINITION AGENCOURT\_6708189 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5749635  
5', mRNA sequence.  
ACCESSION BM919860  
VERSION BM919860.1 GI:19370239  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 976)  
NIH-MGC http://mhc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM12779 row: f column: 04  
High quality sequence stop: 757.  
Location/Qualifiers  
1..976  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5749635"  
/clone\_11b="NIH\_MGC\_120"

```

/1ab_host="DH10B"
/Note="Organ: pooled pancreas and spleen; Vector:
pcmv-Sport6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(boonv) site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC library."
BASE COUNT      285 a      196 c      246 g      248 t      1 others
ORIGIN
Query Match      49.2%; Score 691.8; DB 14; Length 976;
Best local Similarity 87.6%; Pred. No. 1.3e-72;
Matches 832; Conservative 0; Mismatches 18; Indels 100; Gaps 3;
16 GCGGCGGCTGCTGATGTTGTTAGGGAGCAGTAGGGGAAGCTTCTCCCGGA 75
|||||
13 GCGGCGGCTGCTGATTTGTTAGGGAGCAGTAGGGGAAGCTTCTCCCGGA 72
|||||
76 AGAGCTATCTGATTCCTTCTTGATTAACCGTGGCGGAGAGTCAAGGCGCGCT 135
|||||
73 AGAGCTATCTGATTCCTTCTTGATTAACCGTGGCGGAGAGTCAAGGCGCGCT 132
|||||
136 GCGGCGGCTGCTGATGTTGTTAGGGAGCAGTAGGGGAAGCTTCTCCCGGA 195
|||||
133 GCGGCGGCTGCTGATTTGTTAGGGAGCAGTAGGGGAAGCTTCTCCCGGA 192
|||||
196 CCGAATGATTTATTTCAAGTTACTTCTGATTTGCGGAGTCAAGGCTTCTT 255
|||||
193 CCGAATGATTTATTTCAAGTTACTTCTGATTTGCGGAGTCAAGGCTTCTT 252
|||||
256 GCCTTCTTCTGATTTGTTAGGGAGCAGTAGGGGAAGCTTCTCCCGGA 315
|||||
253 GCCTTCTTCTGATTTGTTAGGGAGCAGTAGGGGAAGCTTCTCCCGGA 312
|||||
316 TGGATTTCAAAATTAAGACTATAGAGTTAGAGGGGAAACAACTCAAGCTTCAATA 371
|||||
313 TGGATTTCAAAATTAAGACTATAGAGTTAGAGGGGAAACAACTCAAGCTTCAATA 372
|||||
372 ----- 371
373 ACACAGCAGCGCAGGAAGATTTGCAACATCACTGATTTATACAGAGAGCCCATG 432
372 ----- GAGTCTTCAATATGTTAAAGTGGC 399
433 GCATCATGTTGTTGATGATGTCAGAGATCAGAGTCTTCAATATGTTAAAGTGGC 492
|||||
400 TGCAGAAATATGATGTTGTCAGTGAATGTCACAAATTTGTTAGGGAACAAAT 459
|||||
493 TGCAGAAATATGATGTTGTCAGTGAATGTCACAAATTTGTTAGGGAACAAAT 552
|||||
460 GTGATCTGACCAAGAAAGATGATGATACCAACAGCAGGAATTTGCTGATGCC 519
|||||
553 GTGATCTGACCAAGAAAGATGATGATACCAACAGCAGGAATTTGCTGATGCC 612
|||||
520 TGGAAATTCGTTTGGAAACAGTGTGAATGCAACAGATGAGAAAGCTTTCA 579
|||||
613 TGGAAATTCGTTTGGAAACAGTGTGAATGCAACAGATGAGAAAGCTTTCA 672
|||||
580 TGCAGATGCGCAGTGAATTAAGCAAGTGGTCCGAGCAACAGCTGGTGGCTG 639
|||||
673 TGCAGATGCGCAGTGAATTAAGCAAGTGGTCCGAGCAACAGCTGGTGGCTG 732
|||||
640 AGAAGTCCAAATTTAAATTTAGAGACATCTCCAGTCAAGCAGTGGAGTGGCT 699
|||||
733 AGAAGTCCAAATTTAAATTTAGAGACATCTCCAGTCAAGCAGTGGAGTGGCT 792
|||||
700 AAAATTTGCTCCTATCTTCTCAAGCAATGATTTGCAATGCAATCCCAAGTGA 759
|||||
793 AAAATTTGCTCCTATCTTCTCAAGCAATGATTTGCAATGCAATCCCAAGTGA 852

```

```

OY 760 AACAAATGCTGAATGCTACTGATGCTGACCTAGACAG-ATTCTTACCCTCTC 818
DB 853 AACAAATGCTGAATGCTACTGATGCTGACCTAGACAG-ATTCTTACCCTCTC 912
OY 819 CACAAGGCTCAGAGA--TTGTAAATGCTCAATACGACTTTTAT 865
DB 913 CACAAGGCTCAGAGAATTTAAATGATGATGCTGACCTTTTAT 962

RESULT 13
LOCUS B1086445 837 bp mRNA linear EST 20-JUN-2001
DEFINITION 602849703F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4991233 5',
RNA sequence.
ACCESSION B1086445
VERSION B1086445.1 GI:14504775
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 837)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11008 row: f column: 02
High quality sequence stop: 758.
FEATURES
source location/qualifiers
1 837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4991233"
/clone_lib="NIH_MGC_10"
/cell_line="MGC3"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NCI;
Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5 Kb. Library prepared by Life
Technologies."
BASE COUNT 248 a 134 c 174 g 281 t
ORIGIN
Query Match 49.2%; Score 690.6; DB 13; Length 837;
Best Local Similarity 97.5%; Pred. No. 2.1e-72;
Matches 754; Conservative 0; Mismatches 14; Indels 5; Gaps 5;

OY 557 AACCAATGTAGAACAGTCTTTCATGACGATGACGAGTGAATTAAGCAATGGGTCC 616
DB 1 AACCAATGTAGAACAGTCTTTCATGACGATGACGAGTGAATTAAGCAATGGGTCC 60
OY 617 CGAGAGCAACAGTGTGTGCTGAGAACGCAATGTTAAATTCAGAGCACTCCAGTCA 676
DB 61 CGAGAGCAACAGTGTGTGCTGAGAACGCAATGTTAAATTCAGAGCACTCCAGTCA 120
OY 677 GCATCGAGGAGGAGTGTGCTGATTAATTTGCTTCATCTTTTCACAGCATGAAT 736
DB 121 GCATCGAGGAGGAGTGTGCTGATTAATTTGCTTCATCTTTTCACAGCATGAAT 180
OY 737 TGCATGTGACCAAGTGAAGAAAAAATAATGCTATGCTATGCTATGCTAC 796
DB 181 TGCATGTGACCAAGTGAAGAAAAAATAATGCTATGCTATGCTATGCTAC 240
OY 797 TACAACGATCTCTACGCTCCCAAGAGTCAGATGTAAATGCTATGCTACTT 856

```

```

DB 241 TACAACGATCTCTACGCTCCCAAGAGTCAGATGTAAATGCTATGCTACTT 300
OY 857 TTTTATTTTCCCTGAGTCAAGACAGCTTAATCTTTTTCAGAACTGTTTAAACCTT 916
DB 301 TTTTATTTTCCCTGAGTCAAGACAGCTTAATCTTTTTCAGAACTGTTTAAACCTT 360
OY 917 GTGTGCTGTTTAAATAATGTTGTTAACTGTTGCTTTCGATACCAAGCTGT 976
DB 361 GTGTGCTGTTTAAATAATGTTGTTAACTGTTGCTTTCGATACCAAGCTGT 420
OY 977 TCCCGTGTGTTAGAAATATTTGTTTGTATGTTATTTGGCATGTAGATGCA 1036
DB 421 TCCCGTGTGTTAGAAATATTTGTTTGTATGTTATTTGGCATGTAGATGCA 480
OY 1037 GGTTAGTCTTCTGAAGATGAAGTTCAGCCATTGTTATCAACAGACACACA-GTGTG 1095
DB 481 GGTTAGTCTTCTGAAGATGAAGTTCAGCCATTGTTATCAACAGACACACA-GTGTG 540
OY 1096 TGTCACTTCCATGATCAATAAGTTAGTATGATGATGATGATGATGATGATGATGAT 1155
DB 541 TGTCACTTCCATGATCAATAAGTTAGTATGATGATGATGATGATGATGATGATGATGAT 600
OY 1156 CTTCCTGTGATGATGAT-AAAATGCAAGATTACAG-TATCGATTAATGTTTCTGATA 1213
DB 601 CTTCCTGTGATGATGATTAATAATGCAAGATTACAGTATGATTAATGTTTCTGATA 660
OY 1214 CTTCGATTAATGTTTGTGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1273
DB 661 CTTCGATTAATGTTTGTGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 719
OY 1274 AACCTAATATGTTTAAATAATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAAAT 1326
DB 720 AACCTAATATGTTTAAATAATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAAAT 771

RESULT 14
LOCUS BM706159 689 bp mRNA linear EST 28-FEB-2002
DEFINITION UT-E-DMO-agg-f-10-0-UT-r1 UT-E-DMO Homo sapiens cDNA clone
UT-E-DMO-agg-f-10-0-UT 5', mRNA sequence.
ACCESSION BM706159
VERSION BM706159.1 GI:19019417
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 689)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: mesores@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse
FEATURES
source location/qualifiers
1..689
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="f-10-0-UT"
/clone_lib="UT-E-DMO"

```

```

/tissue_type="lens"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies (T1 phage resistant))"
/Note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-DW0 is a cDNA library containing the following
tissue(s): lens. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GATTTAGCGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
BASE COUNT      195 a      122 c      142 g      230 t
ORIGIN

```

```

Query Match      49.0%; Score 689; DB 14; Length 689;
Best Local Similarity 100.0%; Pred. No. 3.6e-72;
Matches 689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 GGAATTTGCTGATTCCTTGGAAATTCCTTTTGGAAACAGTGTAAAGATGCAACGAA 562
    |||
Db 1 GGAATTTGCTGATTCCTTGGAAATTCCTTTTGGAAACAGTGTAAAGATGCAACGAA 60

QY 563 TGTGAACAGTCTTTCATGAGATGGCAGTGCAGATTTAAAGCAATGGTCCCGGAGC 622
    |||
Db 61 TGTGAACAGTCTTTCATGAGATGGCAGTGCAGATTTAAAGCAATGGTCCCGGAGC 120

QY 623 AACAGCTGTGTGCTGTGAGAGTCCAAATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 682
    |||
Db 121 AACAGCTGTGTGCTGTGAGAGTCCAAATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 180

QY 683 AGTGGAGAGTTGCTGCTAAATTTGGCTCCATCCTTTTTCACAGCAATGAAATTTGCAAT 742
    |||
Db 181 AGTGGAGAGTTGCTGCTAAATTTGGCTCCATCCTTTTTCACAGCAATGAAATTTGCAAT 240

QY 743 CTGAACCCAGTGAAGAAACAAATTCGTAATTCCTGATGCTGATGCTGATGCTGATGCTGAT 802
    |||
Db 241 CTGAACCCAGTGAAGAAACAAATTCGTAATTCCTGATGCTGATGCTGATGCTGATGCTGAT 300

QY 803 AGATTTCTTACCGTCTCCACAAAGTTCAGAGATGTTAAATGTCATATGACTTTTCTT 862
    |||
Db 301 AGATTTCTTACCGTCTCCACAAAGTTCAGAGATGTTAAATGTCATATGACTTTTCTT 360

QY 863 TATTCCTTCTGACTCAAGACAGTCACTTATTCATGACAGTGTAAACCTTTGTGTGC 922
    |||
Db 361 TATTCCTTCTGACTCAAGACAGTCACTTATTCATGACAGTGTAAACCTTTGTGTGC 420

QY 923 TGGTTTAAATTAATGTTGTTTGAATTCCTTGTGCTTCCATGACAGAGTGTCCCT 982
    |||
Db 421 TGGTTTAAATTAATGTTGTTTGAATTCCTTGTGCTTCCATGACAGAGTGTCCCT 480

QY 983 GGTGGTTAGAAATATTTGTTTGAATTTGATATGTCATGTTTATGATGTCAGTTTA 1042
    |||
Db 481 GGTGGTTAGAAATATTTGTTTGAATTTGATATGTCATGTTTATGATGTCAGTTTA 540

QY 1043 GCTCTTCGAAGATTAATGTCAGCAATTTTGTATCAACAGCAAGCAGTGTCTGACT 1102
    |||
Db 541 GCTCTTCGAAGATTAATGTCAGCAATTTTGTATCAACAGCAAGCAGTGTCTGACT 600

QY 1103 TTTCATGATTAATGTTAGTGAAGATTTATATGTAAGTCTGATTTGCTAGTCTTCTCT 1162
    |||
Db 601 TTTCATGATTAATGTTAGTGAAGATTTATATGTAAGTCTGATTTGCTAGTCTTCTCT 660

QY 1163 GTAGAGTTATTAATGGAAGATTTACACTA 1191
    |||
Db 661 GTAGAGTTATTAATGGAAGATTTACACTA 689

```

```

RESULT 15
BI092003      737 bp      mRNA      linear      EST 20-JUN-2001
LOCUS        60285630221 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4997688 5',
DEFINITION   mRNA sequence.
ACCESSION    BI092003
VERSION      BI092003.1 GI:14510333
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 737)
AUTHORS      NIH-MGC http://mhc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs-remail.nih.gov
              Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1025 row: c column: 01
High quality sequence stop: 727.
Location/Qualifiers

```

## FEATURES

```

1..737
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_image="IMAGE:4997688"
    /clone_id="NIH_MGC_10"
    /cell_line="MGC36"
    /lab_host="DH10B"
    /note="Organ: cervix; Vector: pCMV-Sport6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
    Average insert size 1.5 kb. Library prepared by Life
    Technologies."

```

```

BASE COUNT      214 a      135 c      147 g      241 t
ORIGIN

```

```

Query Match      49.0%; Score 688.2; DB 13; Length 737;
Best Local Similarity 97.8%; Pred. No. 4.3e-72;
Matches 719; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

```

```

QY 550 AGAATGCAACGAATGTAGAACAGTCTTTCATGACGATGGCAGTGAATTTAAAGCGAA 609
    |||
Db 1 AGAATGCAACGAATGTAGAACAGTCTTTCATGACGATGGCAGTGAATTTAAAGCGAA 60

QY 610 TGGTCCCGGAGCAACAGCTGGTGTGCTGAGAGTCCAAATGTTAAATTCAGAGCACTC 669
    |||
Db 61 TGGTCCCGGAGCAACAGCTGGTGTGCTGAGAGTCCAAATGTTAAATTCAGAGCACTC 120

QY 670 CAGTCAAGCACTCAGTGTGAGAGTGTGCTGCTAAATTTGCTCCATCCTTTTCTCAGACA 729
    |||
Db 121 CAGTCAAGCACTCAGTGTGAGAGTGTGCTGCTAAATTTGCTCCATCCTTTTCTCAGACA 180

QY 730 ATGAATTTGCAATCGAAGCCCAAGTGAAGAAACAAATTCGCTGAATTTGTCATGTA 789
    |||
Db 181 ATGAATTTGCAATCGAAGCCCAAGTGAAGAAACAAATTCGCTGAATTTGTCATGTA 240

QY 790 GCTCAGTCAACAGATTTCTTACGCTCTCCACAAAGTCAAGATGTTAAATGTCATA 849
    |||
Db 241 GCTCAGTCAACAGATTTCTTACGCTCTCCACAAAGTCAAGATGTTAAATGTCATA 300

QY 850 CTGACTTTTATTTATTCCTTGAAGTCAAGACAGTCACTTATTTTCAGACTGTTT 909
    |||
Db 301 CTGACTTTTATTTATTCCTTGAAGTCAAGACAGTCACTTATTTTCAGACTGTTT 360

QY 910 AACCTTGTGCTGCTGTTTAAATTAATGTTGTTATCCTTGTGCTTCTGATACCA 969
    |||

```



```
Db 361 AACCTTGCTGCTGCTTTATAAATATATGCTGTATACCTTGCTTCCGTATACA 420
QY 970 GACTGTTCCCGGTGGTGTAGATATATTTGTTTGA--TGTATATGSCATGTT 1027
Db 421 GACTGTTCCCGGTGGTGTAGATATATTTGTTTGA--TGTATATGSCATGTT 480
QY 1028 TAGATGTCAGCTTGTAGTCTGTGTAAGATGATGACCATTTGTATCAACACACAA 1087
Db 481 TAGATGTCAGCTTGTAGTCTGTGTAAGATGATGACCATTTGTATCAACACACAA 540
QY 1088 GCAGTGTCTGTGACTTTCATGATGATTAAGTTAGTATGATGATGATGATGATGAT 1147
Db 541 GCAGTGTCTGTGACTTTCATGATGATTAAGTTAGTATGATGATGATGATGATGAT 600
QY 1148 TGTAGTCTTCCCTGTAGAGTATATAATGAA--AGATTACCTATCTGATTAAGTTT 1206
Db 601 TGTAGTCTTCCCTGTAGAGTATATAATGAA--AGATTACCTATCTGATTAAGTTT 660
QY 1207 CTTCATCTCTGCAATATATTTTGGCTGCAGATATGTAATTTGTGCACTATGTA 1266
Db 661 CTTCATCTCTGCAATATATTTTGGCTGCAGATATGTAATTTGTGCACTATGTA 720
QY 1267 ACAAAACACTGAG 1281
Db 721 ACAAAACACTGAG 735
```

Search completed: June 22, 2003, 21:26:05  
Job time : 1910 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 19:45:31 ; Search time 94 Seconds

(Without alignments)  
4583.839 Million cell updates/sec

Title: US-09-820-003B-1

Perfect score: 1405

Sequence: 1 aagcgatgctgctgctgcgcgc.....aaaaaaaaaaaaaaaaaaaaa 1405

Scoring table: IDENTITY\_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2.6/pdata1/lna/5A.COMB.seq:\*\*
- 2: /cgn2.6/pdata1/lna/5B.COMB.seq:\*\*
- 3: /cgn2.6/pdata1/lna/6A.COMB.seq:\*\*
- 4: /cgn2.6/pdata1/lna/6B.COMB.seq:\*\*
- 5: /cgn2.6/pdata1/lna/PCNUS.COMB.seq:\*\*
- 6: /cgn2.6/pdata1/lna/backfiles1.seq:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	13.9	925	2	US-08-916-901-4
2	195	13.9	925	2	US-09-154-602-4
3	171	12.2	171	4	US-09-506-729-59
4	166	11.8	166	4	US-09-506-729-58
5	102.6	7.3	2246	4	US-09-363-708-3
6	102.2	7.3	140	1	US-08-628-417-5
7	102.2	7.3	240	1	US-08-628-417-6
8	102	7.3	1872	4	US-09-801-052-1
9	100.8	7.2	1117	4	US-09-247-3738-33
10	100.4	7.1	1066	1	US-08-157-101A-4
11	100.4	7.1	1813	4	US-09-071-224-3
12	100	7.1	1411	4	US-08-964-127-5
13	100	7.1	1411	4	US-09-496-692-5
14	99.8	7.1	1582	3	US-08-545-196B-10
15	99.8	7.1	1582	3	US-08-545-196B-12
16	99.8	7.1	2447	2	US-09-014-969-14
17	99.4	7.1	1474	1	US-08-702-344-26
18	99.2	7.1	1474	1	US-08-821-994-64
19	99.2	7.1	1683	4	US-09-347-803-11
20	99.2	7.1	3375	4	US-09-370-838-151
21	99	7.0	2671	6	5168051-9
22	98.6	7.0	1798	4	US-09-797-906-1
23	98.4	7.0	2806	4	US-09-653-838-9
24	97.8	7.0	2674	4	US-09-817-180-1
25	97.6	6.9	1447	4	US-09-443-041A-27
26	97.6	6.9	1578	4	US-09-602-877A-95
27	96.8	6.9	1641	1	US-08-300-903A-8

28	95.6	6.8	1098	3	US-09-248-335-35	Sequence 35, Appl
29	95.6	6.8	2323	4	US-09-148-476-24	Sequence 24, Appl
30	95.2	6.8	1882	4	US-09-370-253-1	Sequence 1, Appl
31	95.2	6.8	2634	4	US-09-463-238-3	Sequence 3, Appl
32	95.2	6.8	6871	1	US-08-280-443-1	Sequence 1, Appl
33	95.2	6.8	6871	1	US-08-457-659-1	Sequence 1, Appl
34	95.2	6.8	6871	1	US-08-555-678-1	Sequence 1, Appl
35	95.2	6.8	6871	5	PCT-US95-02275-1	Sequence 1, Appl
36	94.8	6.7	2186	4	US-09-360-545-66	Sequence 66, Appl
37	94.4	6.7	1454	4	US-09-372-422A-19	Sequence 19, Appl
38	94.2	6.7	1733	3	US-09-073-569-1	Sequence 1, Appl
39	93.8	6.7	1051	4	US-09-245-041-10	Sequence 10, Appl
40	93.8	6.7	2184	4	US-08-955-918C-1	Sequence 1, Appl
41	93.8	6.7	2184	4	US-08-697-766A-1	Sequence 1, Appl
42	93.6	6.7	1493	1	US-08-340-820-24	Sequence 24, Appl
43	93.6	6.7	1493	1	US-08-593-535-24	Sequence 24, Appl
44	93	6.6	2082	2	US-08-785-310A-2	Sequence 2, Appl
45	92.8	6.6	1198	3	US-09-248-335-27	Sequence 27, Appl

## ALIGNMENTS

RESULT 1  
US-08-916-901-4  
; Sequence 4, Application US/08916901  
; Patent No. 5892012  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley Neil C.  
APPLICANT: Shah, Purya  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/916,901  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 925 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVERMUT04  
CLONE: 2514506  
US-08-916-901-4  
Query Match 13.9%; Score 195; DB 2; Length 925;  
Best Local Similarity 63.5%; Pred. No. 1e-30; Indels 99; Gaps 2;  
Matches 369; Conservative 0; Mismatches 125;

QY 184 CCAGATGAATCCGGAATATGATTTATTCAGATTACTTGTGATGGCAGCTCAGGG 243  
DB 61 CCGCATGAACCCCGGAATATGATACCTGTTAAAGCTGCTTTTATGGCCAGCTCAGCG 120  
QY 244 TTGGAAGTCTTGGCTTCTTCTTGTAGGATGATATACATATACAGAAAGTATACATCA 303  
DB 121 TGGCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
QY 304 GCACAAATGTTGTGATTTTCAAAATTAAGAACTATAGATTAGAGCGGAAACAAATCAAGC 363  
DB 181 GCACCAATCGGGGTGAGATTCAAGATCCGAACTGAGTGGATGAGCAAACTATCAAC 240  
QY 364 TTCAATA----- 371  
DB 241 TTCAGATCTGGGACACAGCGGGCCAGGAAGCGTTCCGGACCATCTTCCAGTACTACCC 300  
QY 372 -----GAGTCTTCAATATG 387  
DB 301 GGGGGGCTCATGCGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 388 TTAACAGTGGCTCAGAGAAATGATGCTTATGCCAGTAAATGTCAACAAATGTTGG 447  
DB 361 TGAAGCAGTGGCTCAGAGAAATGATGCTTATGCCAGTAAATGTTAGCTCTGG 420  
QY 448 TAGGAACAAATGATCTGACCAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 507  
DB 421 TGGCAACAGAGGAGCTCAGCAACAGAAAGGTTGGACACACCAACCAAGGAGT 480  
QY 508 TTGCTGATTCCTTGAATTCCTTGTGAAACAGTGTAAAGTAAAGTAAAGTAAAGTAAAGT 567  
DB 481 TTGCAAGTCTCTGGGATCCCTTCTTGGAGAGAGGCGCAAGATGCCACCAATGCG 540  
QY 568 AACAGTCTTTCATACAGTGCAGCTGAGATTAAGAAAGGAGTGGTCCGAGCAACAG 627  
DB 541 AGCAGGCTTCATGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 628 CTGCTGCTGCTGAGAGTCCAAATGTTAAATTCAGAGCTCCAGTCAAGAGTCAAGTG 687  
DB 601 CTGGGGG---CGAGCGGCCCAATCTCAAGATCGACAGACCCCTGTAAAGCGGCTGGCG 657  
QY 688 GAGGTGCTGCTA 700  
DB 658 GTGCTGTGCTA 670

RESULT 2  
US-09-154-602-4  
Sequence 4, Application US/09154602  
Patent No. 6300472  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,602  
FILING DATE:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/916,901  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0535  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 925 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVERU04  
CLONE: 2514506  
US-09-154-602-4

Query Match 13.9%; Score 195; DB 4; Length 925;  
Best Local Similarity 63.5%; Pred. No. 1e-30;  
Matches 389; Conservative 0; Mismatches 125; Indels 99; Gaps 2;

QY 184 CCAGATGAATCCGGAATATGATTTATTCAGATTACTTGTGATGGCAGCTCAGGG 243  
DB 61 CCGCATGAACCCCGGAATATGATACCTGTTAAAGCTGCTTTTATGGCCAGCTCAGCG 120  
QY 244 TTGGAAGTCTTGGCTTCTTCTTGTAGGATGATATACATATACAGAAAGTATACATCA 303  
DB 121 TGGCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
QY 304 GCACAAATGTTGTGATTTTCAAAATTAAGAACTATAGATTAGAGCGGAAACAAATCAAGC 363  
DB 181 GCACCAATCGGGGTGAGATTCAAGATCCGAACTGAGTGGATGAGCAAACTATCAAC 240  
QY 364 TTCAATA----- 371  
DB 241 TTCAGATCTGGGACACAGCGGGCCAGGAAGCGTTCCGGACCATCTTCCAGTACTACCC 300  
QY 372 -----GAGTCTTCAATATG 387  
DB 301 GGGGGGCTCATGCGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 388 TTAACAGTGGCTCAGAGAAATGATGCTTATGCCAGTAAATGTCAACAAATGTTGG 447  
DB 361 TGAAGCAGTGGCTCAGAGAAATGATGCTTATGCCAGTAAATGTTAGCTCTGG 420  
QY 448 TAGGAACAAATGATCTGACCAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 507  
DB 421 TGGCAACAGAGGAGCTCAGCAACAGAAAGGTTGGACACACCAACCAAGGAGT 480  
QY 508 TTGCTGATTCCTTGAATTCCTTGTGAAACAGTGTAAAGTAAAGTAAAGTAAAGTAAAGT 567  
DB 481 TTGCAAGTCTCTGGGATCCCTTCTTGGAGAGAGGCGCAAGATGCCACCAATGCG 540  
QY 568 AACAGTCTTTCATACAGTGCAGCTGAGATTAAGAAAGGAGTGGTCCGAGCAACAG 627  
DB 541 AGCAGGCTTCATGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 628 CTGCTGCTGCTGAGAGTCCAAATGTTAAATTCAGAGCTCCAGTCAAGAGTCAAGTG 687  
DB 601 CTGGGGG---CGAGCGGCCCAATCTCAAGATCGACAGACCCCTGTAAAGCGGCTGGCG 657  
QY 688 GAGGTGCTGCTA 700  
DB 658 GTGCTGTGCTA 670

RESULT 3  
US-09-506-729-59  
Sequence 59, Application US/09506729  
Patent No. 6365352



GENERAL INFORMATION:  
APPLICANT: GILLESPIE, DAVID  
TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC  
TITLE OF INVENTION: POLYMERASE CHAIN REACTION  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL  
ADDRESS: DEFENSE COMMAND  
STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)  
CITY: ABERDEEN PROVING GROUND  
STATE: MARYLAND  
COUNTRY: USA  
ZIP: 21010-5423  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/628,417  
APPLICATION NUMBER: US/08/628,417  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BIFFONTI, ULYSSES J  
REGISTRATION NUMBER: 39,908  
REFERENCE/DOCKET NUMBER: DAM 398-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 410-671-1158  
TELEFAX: 410-671-2534  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 140 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: oligodeoxynucleotide  
HYPOTHETICAL: NO  
ANTI-SENSE: YES  
US-08-628-417-5

Query Match 7.3%; Score 102.2; DB 1; Length 140;  
Best Local Similarity 83.5%; Pred. No. 2.6e-12;  
Matches 116; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1267 ACAAAACACTGAGATATGTTTAAATATGTTACTTATGGAAGTAAAAA 1326  
DB 2 AGAGATTAAGACAAATTAACCTTAGAAATAATTTACTTAAAAA 61  
QY 1327 AA 1386  
DB 62 AA 121  
QY 1387 AAAAAAAAAAAAAAAAAAAAA 1405  
DB 122 AAAAAAAAAAAAAAAAAAAAA 140

RESULT 7  
US-08-628-417-6  
Sequence 6, Application US/08628417  
Patent No. 5627054  
GENERAL INFORMATION:  
APPLICANT: GILLESPIE, DAVID  
TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC  
TITLE OF INVENTION: POLYMERASE CHAIN REACTION  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL  
ADDRESS: DEFENSE COMMAND  
STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)  
CITY: ABERDEEN PROVING GROUND  
STATE: MARYLAND  
COUNTRY: USA

ZIP: 21010-5423  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/628,417  
APPLICATION NUMBER: US/08/628,417  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BIFFONTI, ULYSSES J  
REGISTRATION NUMBER: 39,908  
REFERENCE/DOCKET NUMBER: DAM 398-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 410-671-1158  
TELEFAX: 410-671-2534  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: oligodeoxynucleotide  
HYPOTHETICAL: NO  
ANTI-SENSE: YES  
US-08-628-417-6

Query Match 7.3%; Score 102.2; DB 1; Length 240;  
Best Local Similarity 83.5%; Pred. No. 3e-12;  
Matches 116; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1267 ACAAAACACTGAGATATGTTTAAATATGTTACTTATGGAAGTAAAAA 1326  
DB 2 AGAGATTAAGACAAATTAACCTTAGAAATAATTTACTTAAAAA 61  
QY 1327 AA 1386  
DB 62 AA 121  
QY 1387 AAAAAAAAAAAAAAAAAAAAA 1405  
DB 122 AAAAAAAAAAAAAAAAAAAAA 140

RESULT 8  
US-09-801-052-1  
Sequence 1, Application US/09801052  
Patent No. 6368642  
GENERAL INFORMATION:  
APPLICANT: BEASLEY, Ellen  
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE  
FILE REFERENCE: C1001045  
CURRENT APPLICATION NUMBER: US/09/801,052  
CURRENT FILING DATE: 2001-03-08  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1872  
TYPE: DNA  
ORGANISM: Human  
US-09-801-052-1

Query Match 7.3%; Score 102; DB 4; Length 1872;  
Best Local Similarity 80.0%; Pred. No. 5.7e-12;  
Matches 120; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1256 CACACTATGTAACAACAACACTGAAGATATGTTTAAATATGTTACTTATGGAAGTA 1315  
DB 1712 CCACCTTCCAGGCGCAAAAGGCCAGGCTTATATAGTAATAACTGTCTGTAA 1771





```

TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...966
US-08-964-127-5

Query Match 7.1%; Score 100; DB 4; Length 1411;
Best Local Similarity 84.8%; Pred. No. 1.3e-11;
Matches 112; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1273 CAACTGAAGATATGTTTAATTAATATGTACTTAITGGAGCTAAAAA 1332
Db 1270 CAAATAAAAGACATTGTAATAAAGAGACATTTCATAGCCCTGGAAAAA 1329

```

```

Db      1330 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1369
QY      1393 AAAAAAAAAAAAAA 1404
        |||||
Db      1390 AAAAAAAAAAAAAA 1401

RESULT 13
US-09-496-692-5
; Sequence 5, Application US/09496692
; Patent No. 6313271
; GENERAL INFORMATION:
; APPLICANT: Grandearl, Andrew David John
; TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE

```

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/496,692  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/964,127  
FILING DATE: 06-NOV-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Crews, Ph.D., L. Lee  
REGISTRATION NUMBER: P-43,567  
REFERENCE/DOCKET NUMBER: 07334/038001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ. ID NO.: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1411 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear







241 GGGTGGAAAGCTTGCCTCTCTCTAGTTTGCAGATGATACATATACAGAAAGCTTCA 300  
OY 301 TCAGCACAATTTGGTGTGATTTCAAAATAGAACTATAGAGTTAGACGGGAAACATCA 360  
DB 301 TCAGCACAATTTGGTGTGATTTCAAAATAGAACTATAGAGTTAGACGGGAAACATCA 360  
OY 361 AGCTTCAATAGAGCTTCTCAATATATTTAAACAGTGGCTGAGAAATAGATCTTATG 420  
DB 361 AGCTTCAATAGAGCTTCTCAATATATTTAAACAGTGGCTGAGAAATAGATCTTATG 420  
OY 421 CCAGTGAATAATGTCAACAAATTTGTGTAGGGAACAATGTGATGTGACCAACAAGAG 480  
DB 421 CCAGTGAATAATGTCAACAAATTTGTGTAGGGAACAATGTGATGTGACCAACAAGAG 480  
OY 481 TAGTAGACTACACAACAGCAGGAAGAAATTTGCTGATTTCCCTTGAATTCCTTTTGGAA 540  
DB 481 TAGTAGACTACACAACAGCAGGAAGAAATTTGCTGATTTCCCTTGAATTCCTTTTGGAA 540  
OY 541 CCAGTGTCTAAGATTCACAAGATGTAGAACAGTCTTTTATGACATGCGACGTGAGATTA 600  
DB 541 CCAGTGTCTAAGATTCACAAGATGTAGAACAGTCTTTTATGACATGCGACGTGAGATTA 600  
OY 601 AAAAGCAATGGGTCCCGGACCAACAGCTGGTGGTGTGAGAAAGTCAATGTAAATTC 660  
DB 601 AAAAGCAATGGGTCCCGGACCAACAGCTGGTGGTGTGAGAAAGTCAATGTAAATTC 660  
OY 661 AGAGCACTCCAGTCAAGCAGTCAAGTGGAGTGGTCTGCTAAATTTGGCTCCATCTTTT 720  
DB 661 AGAGCACTCCAGTCAAGCAGTCAAGTGGAGTGGTCTGCTAAATTTGGCTCCATCTTTT 720  
OY 721 CTCACAGCAATGAATTTGCATCTCAATCCCAAGTGAAGAAAAAATTTGCTGATTTGTA 780  
DB 721 CTCACAGCAATGAATTTGCATCTCAATCCCAAGTGAAGAAAAAATTTGCTGATTTGTA 780  
OY 781 CTGATAGTAGCTGACACTCAACAGATTTCTACCGCTCCACAAAGGTGAGAGATGTAAA 840  
DB 781 CTGATAGTAGCTGACACTCAACAGATTTCTACCGCTCCACAAAGGTGAGAGATGTAAA 840  
OY 841 TGGTCAATACAGACTTTTCTTTTATCTCCCTGACCTCAAGACAGTCAATCTTATTCAGA 900  
DB 841 TGGTCAATACAGACTTTTCTTTTATCTCCCTGACCTCAAGACAGTCAATCTTATTCAGA 900  
OY 901 ACTGTTTAAACCTTTGTGTGTGTATTAATAATGTGATCTGTTGCTTTC 960  
DB 901 ACTGTTTAAACCTTTGTGTGTGTATTAATAATGTGATCTGTTGCTTTC 960  
OY 961 CTGATACAGACTTTTCCCGTGGTGTGTAATAATTTTGTGTTGATTTTATG 1020  
DB 961 CTGATACAGACTTTTCCCGTGGTGTGTAATAATTTTGTGTTGATTTTATG 1020  
OY 1021 GCATGTTTAAATGTCAAGTGTATGCTTGAAGATGAAGTTCAGCATTTTGTATCAAC 1080  
DB 1021 GCATGTTTAAATGTCAAGTGTATGCTTGAAGATGAAGTTCAGCATTTTGTATCAAC 1080  
OY 1081 AGCAACAGCAATGTCTGCTCACTTCCATGCAATGAAGTTTACTGAGATTTATATGAGA 1140  
DB 1081 AGCAACAGCAATGTCTGCTCACTTCCATGCAATGAAGTTTACTGAGATTTATATGAGA 1140  
OY 1141 TCTGATTTGCTAGTCTTCTCTTGAAGTATTAATGGAAGATTTACACTATTCATTA 1200  
DB 1141 TCTGATTTGCTAGTCTTCTCTTGAAGTATTAATGGAAGATTTACACTATTCATTA 1200  
OY 1201 TAGTTTCTCACTGCTGCAATATATTTGTGGCTGCAAGATTTGTAATTTGTGCACAC 1260  
DB 1201 TAGTTTCTCACTGCTGCAATATATTTGTGGCTGCAAGATTTGTAATTTGTGCACAC 1260  
OY 1261 TATGTAAACAAACAGTGAATATTTTAAATATTTGATCTTATTTGGAAGTAAAAA 1320  
DB 1261 TATGTAAACAAACAGTGAATATTTTAAATATTTGATCTTATTTGGAAGTAAAAA 1320  
OY 1321 AA 1380  
DB 1321 AA 1380

OY 1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405  
DB 1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405  
RESULT 2  
US-09-820-003A-3  
; Sequence 3, Application US/09820003A  
; Patent No. US20020142382A1  
; GENERAL INFORMATION:  
; APPLICANT: MERKULOV, Gennady et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; FILE REFERENCE: CL001196-  
; CURRENT APPLICATION NUMBER: US/09/820,003A  
; CURRENT FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 46050  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (1)...(46050)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-820-003A-3  
Query Match 58.2%; Score 817.2; DB 10; Length 46050;  
Best Local Similarity 99.6%; Pred. No. 2.4e-130;  
Matches 819; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 502 AGGAATTTGCTGATCCCTGGAATTCCTTTTGGAAACCGATGATGAAGATCAACA 561  
DB 43228 AGGAATTTGCTGATCCCTGGAATTCCTTTTGGAAACCGATGATGAAGATCAACA 43297  
OY 562 ATGTAAGACAGTCTTTCATGACGATGGCAGCTGAGATTTAAAGCAGATGGTCCCGAG 621  
DB 43298 ATGTAAGACAGTCTTTCATGACGATGGCAGCTGAGATTTAAAGCAGATGGTCCCGAG 43357  
OY 622 CAACAGCTGGTGGTCTGAGAGATTCATGTTAAATTCACAGACATCCAGTCAAGCAGT 681  
DB 43358 CAACAGCTGGTGGTCTGAGAGATTCATGTTAAATTCACAGACATCCAGTCAAGCAGT 43417  
OY 682 CAGGTGAGGTGGTGGTCTGAAATTTGCTCATGCTCTTTCACAGCAATGAATTTGCA 741  
DB 43418 CAGGTGAGGTGGTGGTCTGAAATTTGCTCATGCTCTTTCACAGCAATGAATTTGCA 43477  
OY 742 TCTGAACCCAGTGAACCAACAAATTCCTGAATTTGATGATGATGATGATGATGATG 801  
DB 43478 TCTGAACCCAGTGAACCAACAAATTCCTGAATTTGATGATGATGATGATGATGATG 43537  
OY 802 CAGATTTCTACCGTCTCCACAAAGTCAAGATTTGAATGTCATATCTGACTTTT 861  
DB 43538 CAGATTTCTACCGTCTCCACAAAGTCAAGATTTGAATGTCATATCTGACTTTT 43597  
OY 862 TTAATCCCTGAGTCAAGACAGTCAATCTTTCGAATTTGCAATGTTTAAACCTTGGTG 921  
DB 43598 TTAATCCCTGAGTCAAGACAGTCAATCTTTCGAATTTGCAATGTTTAAACCTTGGTG 43657  
OY 922 CTGTTTAAATAATATGTGTATATCTTGTCTTCTGATACAGACATGTTCCG 981  
DB 43658 CTGTTTAAATAATATGTGTATATCTTGTCTTCTGATACAGACATGTTCCG 43717  
OY 982 TGTGTGTAGAAATATTTTGTGATGATTTAATGATGATGATGATGATGATGATGATG 1041  
DB 43718 TGTGTGTAGAAATATTTTGTGATGATTTAATGATGATGATGATGATGATGATGATG 43777  
OY 1042 AGTCTTCAAGATGAAGTTCAGCATTTTGTATCAACACAGCAACAGCAGTGTCTGCAC 1101  
DB 43778 AGTCTTCAAGATGAAGTTCAGCATTTTGTATCAACACAGCAGCAGTGTCTGCAC 43837

QY 1102 TTTCATGCAATAAGTTAGTGAATGTTAATGTGAAGATCTGATTTGCTAGTCTTCT 1161  
DB 43838 TTTCATGCAATAAGTTAGTGAATGTTAATGTGAAGATCTGATTTGCTAGTCTTCT 43897  
QY 1162 TGTAGAGTATAAATGGAAGATTTACATCTGATTAATAGTTTCTTCATCTGAT 1221  
DB 43898 TGTAGAGTATAAATGGAAGATTTACATCTGATTAATAGTTTCTTCATCTGAT 43957  
QY 1222 ATAAATTTGGCTGCAGAAATTTGTAAATTTGTCACATCTGATTAACAACAACTGAG 1281  
DB 43958 ATAAATTTGGCTGCAGAAATTTGTAAATTTGTCACATCTGATTAACAACAACTGAG 44017  
QY 1282 ATATGTTAATAAATATGTTACTTATTTGAGAGTAAAAAAA 1323  
DB 44018 ATATGTTAATAAATATGTTACTTATTTGAGAGTAAAAAAA 44059

## RESULT 3

US-09-820-003A-37  
Sequence 37, Application US/09820003A  
Patent No. US20020142382A1  
GENERAL INFORMATION:  
APPLICANT: MERCK/JOY, Genadey, et al.  
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
PROTEINS, AND USES THEREOF  
FILE REFERENCE: C10001196  
CURRENT APPLICATION NUMBER: US/09/820,003A  
CURRENT FILING DATE: 2001-03-29  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 37  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-820-003A-37

Query Match 42.0%; Score 589.8; DB 10; Length 601;  
Best Local Similarity 99.3%; Pred. No. 4,2e-92;  
Matches 591; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 729 AATGAATTTGCAATCTGAACCAAGAGAAAAAAATAGCTGAAATTTGATCTGAT 788  
DB 1 AATGAATTTGCAATCTGAACCAAGAGAAAAAAATAGCTGAAATTTGATCTGAT 60  
QY 789 AGTGCACATCAACAGATTTACCGCTCCCAAGAGTCAAGATTTGAAATGTCAT 848  
DB 61 AGTGCACATCAACAGATTTACCGCTCCCAAGAGTCAAGATTTGAAATGTCAT 120  
QY 849 ACTGACTTTTATATCCCTGAGTCAAGAGTCAAGTCAATTTGATCTGATCTGAT 908  
DB 121 ACTGACTTTTATATCCCTGAGTCAAGAGTCAAGTCAATTTGATCTGATCTGAT 180  
QY 909 AAACCTTTGGTGGCTGTTAATAAATGTTGTAATCTGTTGGCTTCCGATAC 968  
DB 181 AAACCTTTGGTGGCTGTTAATAAATGTTGTAATCTGTTGGCTTCCGATAC 240  
QY 969 AGATGTTCCCGGCTGGTGAATATATTTGTTGATTTGATTTGATTTGATTTGATTT 1028  
DB 241 AGATGTTCCCGGCTGGTGAATATATTTGTTGATTTGATTTGATTTGATTTGATTT 300  
QY 1029 AGATGTTCCCGGCTGGTGAATATATTTGTTGATTTGATTTGATTTGATTTGATTT 1088  
DB 301 AGATGTTCCCGGCTGGTGAATATATTTGTTGATTTGATTTGATTTGATTTGATTT 360  
QY 1089 CAGTGTCTGCTCCTTCATGATCAAGATTTAGTGAATTTGATTTGATTTGATTTGATTT 1148  
DB 361 CAGTGTCTGCTCCTTCATGATCAAGATTTAGTGAATTTGATTTGATTTGATTTGATTT 420  
QY 1149 GCTAGTCTTCTCTTGTAGAGTTAATGGAAGATTTACATCTGATTTGATTTGATTT 1208  
DB 421 GCTAGTCTTCTCTTGTAGAGTTAATGGAAGATTTACATCTGATTTGATTTGATTTGATTT 480

QY 1209 TCATCTCTGCATATAATTTGTCGTCAGAAATATGTGAATTTGTCACACTATGTAC 1268  
DB 481 TCATCTCTGCATATAATTTGTCGTCAGAAATATGTGAATTTGTCACACTATGTAC 540  
QY 1269 AAAAAGTGTGAATTTGTTAATATTTGTTGATTTGTTGATTTGTTGATTTGTTGATTT 1323  
DB 541 AAAAAGTGTGAATTTGTTAATATTTGTTGATTTGTTGATTTGTTGATTTGTTGATTT 595

## RESULT 4

US-09-920-300A-374/C  
Sequence 374, Application US/09920300A  
Patent No. US20020136728A1  
GENERAL INFORMATION:  
APPLICANT: King, Gordon E.  
APPLICANT: Meagher, Madeleine Joy  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
OF INHERITED AND ACQUIRED COLON CANCER  
FILE REFERENCE: 210121.547  
CURRENT APPLICATION NUMBER: US/09/920,300A  
CURRENT FILING DATE: 2001-07-31  
NUMBER OF SEQ ID NOS: 1789  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 374  
LENGTH: 506  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-920-300A-374

Query Match 36.0%; Score 506; DB 10; Length 506;  
Best Local Similarity 100.0%; Pred. No. 8,7e-78;  
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 770 CCTGAATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829  
DB 506 CCTGAATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447  
QY 830 GAGATTTGAAATGCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 889  
DB 446 GAGATTTGAAATGCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 387  
QY 890 TCATTTTCAAGAGCTTTTAAACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949  
DB 386 TCATTTTCAAGAGCTTTTAAACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327  
QY 950 TGTGCTTCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1009  
DB 326 TGTGCTTCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267  
QY 1010 TGTGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069  
DB 266 TGTGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 207  
QY 1070 TGTGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129  
DB 206 TGTGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 147  
QY 1130 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189  
DB 146 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 87  
QY 1190 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249  
DB 86 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 27  
QY 1250 TGTGTCACACTATGTAACAAACA 1275  
DB 26 TGTGTCACACTATGTAACAAACA 1

```
RESULT 5
US-10-033-528-374/c
; Sequence 374, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033.528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 374
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-374
```

```
Query Match 36.0%; Score 506; DB 12; Length 506;
Best Local Similarity 100.0%; Pred. No. 8,7e-78;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 770 CCTGAATGTACTGTAGTACGCTGACATACAAAGATTCTTACGCTCCACAAAGTCA 829
DB 506 CCTGAATGTACTGTAGTACGCTGACATACAAAGATTCTTACGCTCCACAAAGTCA 447
OY 830 GAGATGTAAATGGCAATGACTGACTTTTATTCCTTCTGACTCAAGACGCTACT 889
DB 446 GAGATGTAAATGGCAATGACTGACTTTTATTCCTTCTGACTCAAGACGCTACT 387
OY 890 TCATTTTCAGAACTGTTTAAACCTTGTGCTGCTTATATAAATATGTTGATATCC 949
DB 386 TCATTTTCAGAACTGTTTAAACCTTGTGCTGCTTATATAAATATGTTGATATCC 327
OY 950 TTGTTGCTTCTGATACCAAGCTGTTCCGCGTGGTGGTAGAATATATTTTGTGGA 1009
DB 326 TTGTTGCTTCTGATACCAAGCTGTTCCGCGTGGTGGTAGAATATATTTTGTGGA 267
OY 1010 TGTATATATGGCAGTGTAGATGCAAGCTTATGCTTCTTCAAGATGATGACCCATT 1069
DB 266 TGTATATATGGCAGTGTAGATGCAAGCTTATGCTTCTTCAAGATGATGACCCATT 207
OY 1070 TTGATATCAACAGACAAAGAGTGTCTGTCACTTTCATGCAATTAAGTTAGTGATGT 1129
DB 206 TTGATATCAACAGACAAAGAGTGTCTGTCACTTTCATGCAATTAAGTTAGTGATGT 147
OY 1130 TATATGTAGATGCTGATTTGCTTCTTCTTGTAGATTATAAATGGAAGATTACAC 1189
DB 146 TATATGTAGATGCTGATTTGCTTCTTCTTGTAGATTATAAATGGAAGATTACAC 87
OY 1190 TATGATATATATTTCTTCTTCTTCTGCAATATATTTGGCGTGCAGATATTTGAT 1249
DB 86 TATGATATATATTTCTTCTTCTTCTGCAATATATTTGGCGTGCAGATATTTGAT 27
OY 1250 TTGTTGCACACTATGTAACAAACA 1275
DB 26 TTGTTGCACACTATGTAACAAACA 1
```

```
RESULT 6
US-09-864-761-1742/c
; Sequence 1742, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```

```
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmca-x-1
; CURRENT APPLICATION NUMBER: US/09/864.761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1742
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007318.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
US-09-864-761-1742
```

```
Query Match 33.2%; Score 467; DB 10; Length 487;
Best Local Similarity 100.0%; Pred. No. 4.1e-71;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 502 AGAATTTGCTGATCCCTTGGAATTCGTTTGGAAACCAAGGCTAAGAATGCAACA 561
DB 467 AGAATTTGCTGATCCCTTGGAATTCGTTTGGAAACCAAGGCTAAGAATGCAACA 408
OY 562 ATGTAGAACAGCTCTTCATGACGATGCGACGTGATTAATAAAGCGAATGGTCCGAG 621
DB 407 ATGTAGAACAGCTCTTCATGACGATGCGACGTGATTAATAAAGCGAATGGTCCGAG 346
OY 622 CAACAGCTGGTGTCTGAGAACATGTCANATGTTAAATTCAGAGCACTCCAGTCAAG 681
```





```

; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-040-862-4644

```

Query Match 26.1%; Score 367; DB 9; Length 367;  
Best Local Similarity 100.0%; Pred. No. 4.9e-54;  
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	909	AAACCTTGTCGGGTTATAAATATGTGTATCCCTTGCTTCCTCATAC	968
Db	1	AAACCTTGTCGGGTTATAAATATGTGTATCCCTTGCTTCCTCATAC	60
QY	969	AGACTGTTCCCGTGTTGGTTAGAAATATATTTGGTTGATGTTATATGGACGTTT	1028
Db	61	AGACTGTTCCCGTGTTGGTTAGAAATATATTTGGTTGATGTTATATGGACGTTT	120
QY	1029	AGATGTCAAGTTTAGTCTCTGAAAGATGAATTCAGCCATTTTGTATATAACAGACAAAG	1088
Db	121	AGATGTCAAGTTTAGTCTCTGAAAGATGAATTCAGCCATTTTGTATATAACAGACAAAG	180
QY	1089	CAGGTGTCGTACATTCATCAATAAAGTTAGTACATGTTATATGTAATCAATCGATTT	1148
Db	181	CAGGTGTCGTACATTCATCAATAAAGTTAGTACATGTTATATGTAATCAATCGATTT	240
QY	1149	GCTAGTCTTCCTGTAGAGATTATAAATGAAAGATTACACTATCTGATTAATAGTTCT	1208
Db	241	GCTAGTCTTCCTGTAGAGATTATAAATGAAAGATTACACTATCTGATTAATAGTTCT	300
QY	1209	TCATACCTCGCATTAATTTGTGGCTGCAGATAATGTAATTTGTGCACACTATGTAAC	1268
Db	301	TCATACCTCGCATTAATTTGTGGCTGCAGATAATGTAATTTGTGCACACTATGTAAC	360
QY	1269	AAAACAA 1275	
Db	361	AAAAACA 367	

```

RESULT 9
US-09-918-995-4909
: Sequence 4909, Application US/09918995
: Publication No. US20030073623A1
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: FROM SEVERAL CDNA LIBRARIES
: FILE REFERENCE: 20411-756
: CURRENT APPLICATION NUMBER: US/09/918, 995
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: US/09/235, 076
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ ID NOS: 38054
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4909
: LENGTH: 436
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(436)
: OTHER INFORMATION: n = A,T,C or G
US-09-918-995-4909

```

Query Match	24.3%	Score 341.8	DB 9	Length 436
Best Local Similarity	98.9%	Pred. No. 1.1e-49		
Matches	365	Conservative	0	Mismatches 2
				Indels 2
				Gaps 2
QY	3	GCAGTAGCTGAGT-GCGCGGCTGCTGATTTGTCTTAGGGGAGCGAGTAGGGGAAGACG	61	
Db	69	GCAGTAGCTGAGTGGCGGGCTGCTGATTTGTCTTAGGGGAGCGAGTAGGGGAAGACG	128	
QY	62	TTTCTCTCCCGGAACCTCATCTCATTTCTTTTCATTACCGGTGGCGGGAGAG	121	

Db	129	TTTGTCTCCCGAAGACGCTATCTCATTCCTTTCTTTCGATTACCGGTGGCGGGAG	188
Oy	122	TCAGGGCGGGGGTGTGGCGAAGGGCGGGCGGGCGGCACTGCATGCAT	181
		TT	
Db	189	TACAGGGGGGGGTGTGGCGCAAGGGCGGGGTGGCGGGCGCACTGCATGCAT	248
Oy	182	GTCCAGCATGAATCCGCAATATGATTATTATTATTCAGTTACTTCGATTGGCGACTAGG	241
		TT	
Db	249	GTCCACATGAATCCGCAATATGATTATTATTATTCAGTTACTTCGATTGGCGACTAGG	308
Oy	242	GGTGGAAAGTCTGGCTCTCTCTAGGTTGGCAATGATATCATATACAGAAACTCAT	301
		TT	
Db	309	GGTGGAAAGTCTGGCTCTCTCTAGGTTGGCAATGATATCATATACAGAAACTCAT	368
Oy	302	CAGCACAATTGGTGGATTTCAAAATAAGACTATAGATTAGACGGGAAACCAATCAA	361
		TT	
Db	369	CACACAAATTGGTGGATGATGC - TAATAAGAACTATAGATTAGACGGGAAACCAATCAA	427
Oy	362	GCTTCAAT 370	
		TTTTTTTT	
Db	428	GCTTCAAT 436	

```

RESULT 10
US-09-960-352-5427
: Sequence 5427, Application US/09960352
: Patent No. US20020137139A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengbing
: APPLICANT: Byatt, John C.
: APPLICANT: Mathialagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
: FILE REFERENCE: 16511.006/37-21(10298)C
: CURRENT APPLICATION NUMBER: US/09/960.352
: CURRENT FILING DATE: 2001-09-24
: NUMBER OF SEQ ID NOS: 15112
: SEQ ID NO 5427
: LENGTH: 353
: TYPE: DNA
: ORGANISM: Bos taurus
: OTHER INFORMATION: Clone ID: 23-LIB34-061-Q1-E1-F3
US-09-960-352-5427

```

	Query Match	Best Local Similarity	24.0%;	Score 337.2;	DB 10;	Length 353;	
	Matches	350;	Conservative	98.9%;	Pred. No. 6e-49;	Mismatches 0;	Indels 1;
							Gaps 1;
QY	946	ATCCCTGTGCTTCCGTGATACGACAGCTGTTTCCCGTGTTGTTAGAAATATATTTCCT					1005
Db	1	ATCTTTGGTGGCTTCCGTATACGACAGCTGTTTCC-GTGGTGGTTAGAAATATATTTCCT					59
QY	1006	TTGATGTTTAAATTTGGCATGTTTATGATGTCAAGTTTAACTCTTCTGAAAGATGAAGTTCAC					1065
Db	60	TTGATGTTTAAATTTGGCATGTTTAAATGATGTCAGTTTAACTCTTCTGAAAGATGAAGTTCAC					119
QY	1066	CATTTTGTATCAAAACAGCACAAGAGAGTTCGTCACTTCATTCATGCATATGAAGTTTAGTAG					1125
Db	120	CATTTTGTATCAAAACAGCACAAGAGAGTTCGTCACTTCATTCATGCATATGAAGTTTAGTAG					179
QY	1126	ATGTTAATATGATGATCTGATTTCTGATTCCTCTCTAGAGTTAATTAATGAAAGATT					1185
Db	180	ATGTTAATATGATGATCTGATTTCTGATTCCTCTCTAGAGTTAATTAATGAAAGATT					239
QY	1186	ACACTATCTGATTAATAGTTTCTTCATACCTGTCATATTAATTTGGCTGCAGAAATATG					1245
Db	240	ACACTATCTGATTAATAGTTTCTTCATACCTGTCATATTAATTTGGCTGCAGAAATATG					299
QY	1246	TAAATTTGTGACACTATGTAACAAACAACGTGAAGATATCTTAAATTAATT					1299
Db	300	TAAATTTGTGACACTATGTAACAAACAACGTGAAGATATCTTAAATTAATT					353



0Y	228	ATTGGGACACTCAGGGGTTGGAAAGCTTGCCCTTCCTTAGGTTGGAGATGATACATAT	287
	200		
	200	ATTGGGACACTCAGGGGTTGGAAAGCTTGCCCTTCCTTAGGTTGGAGATGATACATAT	229
0Y	288	ACAGAAAGCTACATCAGCACAATTTGGTGGATTTCAAAATATAGAACTATGAGTTAGAC	347
	260	ACAGAAAGCTACATCAGCACAATTTGGTGGATTTCAAAATATAGAACTATGAGTTAGAC	319
0Y	348	GGGAAATCAATCAAGCTTCAATA	371
	320	GGGAAATCAATCAAGCTTCAATA	343

```

1          RESULT 14
2  US-09-294-093B-4584
3  Sequence 4584, Application US/09294093B
4  Patent No. US20010051335A1
5  GENERAL INFORMATION:
6  APPLICANT:          Lalagudi, Raghunath, V.
7  APPLICANT:          Ito, Laura, Y.
8  APPLICANT:          Sherman, Bradley, K.
9  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
10 FILE REFERENCE:     PL-0009 US
11 CURRENT APPLICATION NUMBER: US/09/294,093B
12 CURRENT FILING DATE: 1999-04-16
13 PRIOR APPLICATION NUMBER: 60/082,567
14 PRIOR FILING DATE:   April 21, 1998
15 NUMBER OF SEQ ID NOS: 6207
16 SOFTWARE:            PERL Program
17 SEQ ID NO 4584
18 LENGTH: 290
19 TYPE: DNA
20 ORGANISM: Zea mays
21 FEATURE:
22 NAME/KEY: misc_feature
23 OTHER INFORMATION: Incyte ID No. US20010051335A1 700354602H1
24 NAME/KEY: unsure
25 LOCATION: 2, 88, 216, 219, 247, 283
26 OTHER INFORMATION: a, t, c, g, or other
27 US-09-294-093B-4584

```

	Query Match	Best Local Similarity	18.6%;	Score 261;	DB 10;	Length 290;
	Matches	Conservative	97.3%;	Pred. No. 6, 3e-36;	Mismatches 6;	Indels 2; Gaps 2
QY	540	ACCGAGCTTAAGATGTCACAGCAAGTAGACAGTCTTTTCATGACGATGGCAGCTAGATT				599
Db	1	AMCAGGTGTAAGAAATGCACAACAAATGTAAACAGTCTTTCATGACCATGGCAGCTAGATT				60
QY	600	AAAAAGCAATGGGATCCCGAGACAACAGCTGGTGT -GCTGAGAGTCCAAATGTAAAT				658
Db	61	AAAAAGGAATGGGATCCCGAGACAACACTGGTGGCTGAGAGTCCAAATGTAAAT				120
QY	659	TCACAGCACTCCAGTCAAGCAAGTCAAGGTGGAGGTTGCTGCTAAATTTTGCTCCATCCTT				718
Db	121	TCACAGCACTCCAGTCAAGCAAGTCAAGGTGGAGGTTGCTGCTAAATTTTGCTCCATCCTT				180
QY	719	TTCTCAGCAGCAATGAATTTGCAATCTGAAACCCAAAGTGAACAAAACAAATTTGGCTAATTTG				778
Db	181	TTCTCAGCAGCAATGAATTTGCAATCTGAAACCCAAAGTGAACAAAACAAATTTGGCTAATTTG				240
QY	779	TACGTGATGTAGCTGCACACTACACAGATTTCTTACCGTCTCCACAAAGTCA				829
Db	241	TACGTGATGTAGC -GCACCTACACAGATTTCTTACCGTCTCCACAAAGTCA				290

RESULT 15  
US-10-079-623-231  
; Sequence 231, Application US/10079623  
; Patent No. US20020169302A1  
; GENERAL INFORMATION:  
; APPLICANT: Havukkala, Ilkka J.

```

1  APPLICANT: Glenn, Matthew
2  APPLICANT: Grigor, Murray R.
3  APPLICANT: Molenaar, Adrian J.
4  TITLE OF INVENTION: Compositions isolated from bovine
5  TITLE OF INVENTION: mammary gland and methods for their use.
6  FILE REFERENCE: 11000.1044c3
7  CURRENT APPLICATION NUMBER: US/10/079,623
8  CURRENT FILING DATE: 2002-02-19
9  NUMBER OF SEQ ID NOS: 370
10 SOFTWARE: FastSeq for Windows Version 4.0
11 SEQ ID NO 231
12     LENGTH: 297
13     TYPE: DNA
14     ORGANISM: Bovine
15 US-10-079-623-231

```

	Query Match	18.5%;	Score 260.7; DB 9;	Length 297;
	Best Local Similarity	94.6%;	Pred. No. 7.4e-36;	
	Matches	281;	Conservative	0; Mismatches 14; Indels 2; Gaps 1;
QY	16	GCGGCGCTGCTGATTGTCTTACGAGCAGGATAGGG--AAGACTTTGCTCTCCG	73	
Db	1	GTGGCGGTGCTGATTGTCTTACGAGGAGGAGTGAGGATTAAGACCTTTGCTCTCTG	60	
QY	74	GAAGAGCCTATCTATCTCTTCTTTCGATTACCGCTGGCCGCGAGATCAGGCGCGG	133	
Db	61	GAACAGCCTATCTATCTCTTCTTCTGACTACCGCTGGCGCGAGATCAGGCGCGG	120	
QY	134	CTGGCGCAGCAAGGGCGGCGGTGGCGCGCGCGCAGCTGACATGTCACATGATA	193	
Db	121	CTGACGACGACAGGGCGGCGGTGGCGCGCGCGCGCTGACATGTCACATGATA	180	
QY	194	TCCCGAATATGATTATTTATTCAGTTACTCTGATTGGCACACTAGGGGTTGAAATC	253	
Db	181	TCCCGAATATGATTATTTATTCAGTTACTCTGATTGGCACACTGAGGGTTGAAATC	240	
QY	254	TTGGCTCTCTGTTAGGTTTGAGATGATACATATACAGAAAGCTACATCACCACAT	310	
Db	241	TTGGCTCTCTGTTAGGTTTGAGATGATACATATACAGAAAGCTACATCACCACAT	297	

Search completed: June 22, 2003, 22:38:01  
Job time : 228 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:15:24 ; Search time 70 seconds  
(Without alignments)  
329.319 Million cell updates/sec

Title: US-09-820-003b-2

Perfect score: 886  
Sequence: 1 MSNNPEYDYLFLKLLIGDS.....EKSNVRIQSTPYKSGGGCC 173

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: A\_Geneseq\_101002.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	860	97.1	205	21	Gene 44 human secr
2	860	97.1	205	21	Human secreted pro
3	817	92.2	198	21	Human secreted pro
4	768.5	86.7	221	22	Human polypeptide
5	768.5	86.7	224	21	Breast and ovarian
6	766.5	86.5	201	20	Human Rab protein
7	766.5	86.5	201	22	Novel human secret
8	683	77.1	205	22	Drosophila melanog
9	637	71.9	202	21	Arabidopsis thaliana
10	633	71.4	202	21	Arabidopsis thaliana

11	619.5	69.9	258	21	AA030498	Arabidopsis thaliana
12	618.5	69.8	203	21	AA030499	Arabidopsis thaliana
13	612	69.1	218	21	AA030710	Arabidopsis thaliana
14	612	69.1	221	21	AA030709	Arabidopsis thaliana
15	612	69.1	254	21	AA045323	Arabidopsis thaliana
16	580	65.5	197	21	AA048820	Arabidopsis thaliana
17	580	65.5	197	21	AA048820	Arabidopsis thaliana
18	442	50.1	218	21	AA019220	Arabidopsis thaliana
19	442	49.9	216	21	AA080006	Arabidopsis thaliana
20	429.5	48.5	215	21	AA035215	Arabidopsis thaliana
21	429.5	48.5	224	21	AA047826	Arabidopsis thaliana
22	428.5	48.4	216	21	AA080688	Arabidopsis thaliana
23	428.5	48.4	253	21	AA039445	Arabidopsis thaliana
24	428.5	48.4	253	21	AA039445	Arabidopsis thaliana
25	428.5	48.4	254	21	AA080687	Arabidopsis thaliana
26	419	47.3	223	22	AB041333	Human rab8 homolog
27	417	47.1	221	23	AB041333	Human rab8 homolog
28	414	46.7	207	22	AA067154	Amino acid sequenc
29	414	46.7	207	22	AA067154	Human protein sequ
30	409.5	46.2	335	22	AA023366	Novel human diagno
31	408	46.0	207	22	AB071647	Drosophila melanog
32	405	45.7	163	21	AA018482	Zea mays protein f
33	403.5	45.5	200	21	AA019165	Amino acid sequenc
34	403.5	45.5	200	21	AA019165	Human protein sequ
35	403.5	45.5	200	22	AA059340	Human prostate can
36	403.5	45.5	218	21	AA056933	Human protein sequ
37	403	45.5	130	21	AA035036	Arabidopsis thaliana
38	400	45.1	141	21	AA018483	Novel human diagno
39	399.5	44.1	218	22	AB07266	Novel human diagno
40	393	44.4	201	21	AA099982	Canine Rab10 prote
41	393	44.4	204	22	AB070670	Drosophila melanog
42	390.5	44.1	209	22	AB023365	Novel human diagno
43	390.5	44.1	246	21	AA05196	Lung cancer associ
44	367	43.7	199	21	AA05980	Human Rab10 prote
45	367	43.7	199	21	AA05980	Human Rab10 prote

ALIGNMENTS

RESULT 1  
AAB34843 standard; Protein, 205 AA.  
AAB34843:  
26-JAN-2001 (first entry)  
Gene 44 human secreted protein homologous amino acid sequence #131.  
Human; secreted protein; diagnosis; cytosolic; immunosuppressive;  
neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;  
antibacterial; antifungal; antiparasitic; cardiant; gene therapy;  
cancer; immune disorder; cardiovascular disorder; wound healing;  
neurological disease; infectious disease; chromosome identification.  
Homo sapiens.  
WO200058356-A1.  
05-OCT-2000.  
22-MAR-2000; 2000WO-US07535.  
26-MAR-1999; 99US-0126511.  
17-DEC-1999; 99US-0172413.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Ruben SM, Komatsoulis G;  
WPI: 2000-594639/56.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful  
PT in the prevention, treatment and diagnosis of cancer, immune disorders,  
PT cardiovascular disorders and neurological diseases -  
XX  
PS Disclosure; Page 415-416; 425pp; English.

CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the  
CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to  
CC AAB34852 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissues  
CC and cells the genes are expressed in. Examples of activities include:  
CC cytosolic; immunosuppressive; neurotropic; neuroprotective; antiviral;  
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
CC vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic; and  
CC cardiant. The polynucleotides and polypeptides are useful for  
CC preventing, treating or ameliorating a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. The polynucleotides are  
CC useful for chromosome identification. They are also useful as probes for  
CC diagnosing a disorder related to the female reproductive system,  
CC particularly breast and/or ovarian cancer. They are also useful in the gene  
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,  
CC agonists and antagonists from the present invention are useful in the  
CC diagnosis, treatment and prevention of cancer, immune disorders,  
CC cardiovascular disorders, wound healing, neurological diseases and  
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence  
CC used in the exemplification of the present invention.

XX Sequence 205 AA;

SO Query Match 97.1%; Score 860; DB 21; Length 205;

Best Local Similarity 84.4%; Pred. No. 7.6e-79;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLKLLIGDSGVGKSCLLRRADPTYSYSTIGVDFKRTIELDGTI 60  
DB 1 MSSNNPEYDYLKLLIGDSGVGKSCLLRRADPTYSYSTIGVDFKRTIELDGTI 60  
QY 61 KIQI-----ESFNVKOMLQEIDRYASENNVKL 88  
DB 61 KIQIWDAGORFRITTSYRGAGIIVYDVTDOESFNVKOMLQEIDRYASENNVKL 120  
QY 89 LVGNKCDLTKKVVDTTAKFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGA 148  
DB 121 LVGNKCDLTKKVVDTTAKFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGA 180  
QY 149 TAGGAESKNVKIOSTPVKQSGGCC 173  
DB 181 TAGGAESKNVKIOSTPVKQSGGCC 205

Db

RESULT 2  
AAB34844 standard; Protein: 205 AA.

XX AAB34844;

XX 26-JAN-2001 (first entry)

XX Human secreted protein sequence encoded by gene 44 SEQ ID NO:132.

XX Human; secreted protein; diagnosis; cytosolic; immunosuppressive;  
XX neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
XX antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;  
XX antibacterial; antifungal; antiparasitic; cardiant; gene therapy;  
XX cancer; immune disorder; cardiovascular disorder; wound healing;  
XX neurological disease; infectious disease; chromosome identification.  
XX Homo sapiens.

PN WO200058356-A1.

PD 05-OCT-2000.

XX 22-MAR-2000; 2000WO-US07535.

XX 26-MAR-1999; 99US-0126511.

XX 17-DEC-1999; 99US-0172413.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruden SM, Komatsoulis G;

XX WPI: 2000-594639/56.

XX Disclosure; Page 416-417; 425pp; English.

CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the  
CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to  
CC AAB34852 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissues  
CC and cells the genes are expressed in. Examples of activities include:  
CC cytosolic; immunosuppressive; neurotropic; neuroprotective; antiviral;  
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
CC vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic; and  
CC cardiant. The polynucleotides and polypeptides are useful for  
CC preventing, treating or ameliorating a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. The polynucleotides are  
CC useful for chromosome identification. They are also useful as probes for  
CC diagnosing a disorder related to the female reproductive system,  
CC particularly breast and/or ovarian cancer. They are also useful in the gene  
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,  
CC agonists and antagonists from the present invention are useful in the  
CC diagnosis, treatment and prevention of cancer, immune disorders,  
CC cardiovascular disorders, wound healing, neurological diseases and  
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence  
CC used in the exemplification of the present invention.

XX Sequence 205 AA;

SO Query Match 97.1%; Score 860; DB 21; Length 205;

Best Local Similarity 84.4%; Pred. No. 7.6e-79;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLKLLIGDSGVGKSCLLRRADPTYSYSTIGVDFKRTIELDGTI 60  
DB 1 MSSNNPEYDYLKLLIGDSGVGKSCLLRRADPTYSYSTIGVDFKRTIELDGTI 60  
QY 61 KIQI-----ESFNVKOMLQEIDRYASENNVKL 88  
DB 61 KIQIWDAGORFRITTSYRGAGIIVYDVTDOESFNVKOMLQEIDRYASENNVKL 120  
QY 89 LVGNKCDLTKKVVDTTAKFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGA 148  
DB 121 LVGNKCDLTKKVVDTTAKFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGA 180  
QY 149 TAGGAESKNVKIOSTPVKQSGGCC 173  
DB 181 TAGGAESKNVKIOSTPVKQSGGCC 205

Db

RESULT 3  
AAB34816 standard; Protein: 198 AA.

XX AAB34816;

Dt	26-JAN-2001	(first entry)
Xx	Human secreted protein sequence encoded by gene 44 SEQ ID NO:104.	
Dd		
Xx	Human; secreted protein; diagnosis; cytotaxic; immunosuppressive;	
Rv	nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;	
Rv	antidiabetic; antiinflammatory; antitumor; vulnery; anticovulant;	
Rv	antibacterial; antifungal; antiparasitic; cardine; gene therapy;	
Rv	cancer; immune disorder; cardiovascular disorder; wound healing;	
Rv	neurological disease; infectious disease; chromosome identification.	
Xx		
Xx	Homo sapiens.	
Oo		
Fm	WO200058356-A1.	
Pd		
Pd	05-OCT-2000.	
Pf		
Pf	22-MAR-2000; 2000WO-US07535.	
Xx		
Xx	26-MAR-1999; 99US-0126511.	
Px	17-DEC-1999; 99US-0172413.	
Pa		
Pa	(HUMA-) HUMAN GENOME SCI INC.	
Pi		
Pi	Rosen CA, Ruben SM, Komatsoulis G;	
Xx		
Xx	WPI: 2000-594639/56.	
Dr		
Dr	N-PSDB; AAC60009.	
Pt		
Pt	Fifty nucleotide acid molecules encoding human secreted proteins, useful	
Pt	in the prevention, treatment and diagnosis of cancer, immune disorders,	
Pt	cardiovascular disorders and neurological diseases -	
Xx		
Ps	Claim 1; Page 393-394; 425pp: English.	
Cc	The polynucleotide sequences given in AAC59966 to AAC60015 encode the	
Cc	human secreted proteins given in AAB34773 to AAB34822. AAB34823 to	
Cc	AAB34832 represent human secreted polypeptide sequences and proteins	
Cc	homologous to them, which are given in the exemplification of the present	
Cc	invention. Human secreted proteins have activities based on the tissues	
Cc	and cells the genes are expressed in. Examples of activities include:	
Cc	cytotoxic; immunosuppressive; nootropic; neuroprotective; antiviral;	
Cc	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;	
Cc	vulnerary; anticovulant; antibacterial; antitumoral; antiparasitic; and	
Cc	cardiant. The polynucleotides and polypeptides are useful for	
Cc	preventing, treating or ameliorating a medical condition in e.g. humans,	
Cc	mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The	
Cc	polypeptides can also be used as a food additive or preservative to	
Cc	increase or decrease storage capabilities. The polynucleotides are	
Cc	useful for chromosome identification. They are also useful as probes for	
Cc	diagnosing a disorder related to the female reproductive system,	
Cc	particularly breast and/or ovarian cancer. They are also useful in the	
Cc	therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,	
Cc	agonists and antagonists from the present invention are useful in the	
Cc	diagnosis, treatment and prevention of cancer, immune disorders,	
Cc	cardiovascular disorders, wound healing, neurological diseases and	
Cc	infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence	
Cc	used in the exemplification of the present invention.	
Xx		
Xx		
Sq	Sequence 198 AA:	
	Query Match 92.28; Score 817; Db 21; Length 198;	
	Best Local Similarity 83.86; Fred. No. 1.6e-74;	
	Matches 165; Conservative 0; Mismatches 0; Indels 32; Gaps 1	
Oy	9 DYLFKLLIGDSGVKSCLLLRPADYTBESYSTGTGDFKRTIELDKTKTIQI--- 64	
Db	1 DYLFKLLIGDSGVKSCLLLRPADYTBESYSTGTGDFKRTIELDKTKTIKIIMPTA 60	
Oy	65 -----ESPNNVKKOMLOEIDRKASBNVKLLVGNCKDL 96	
Db	61 GORFTITSYSGAHGIIVYDWTDESNVNVKKOMLOEIDRKASBNVKLLVGNCKDL 120	

OY	97	TTKKVVDYTTAKKEPADSLGPIPLETSANNTNVEOSFMTMAETKTRMGPATGAGERS	15
DB	121	TTKKVVDYTTAKKEPADSLGPIPLETSANNTNVEOSFMTMAETKTRMGPATGAGERS	180
OY	157	NVKIOSTPYKSGGGCC	173
DB	181	NVKIOSTPYKSGGGCC	197
RESULT 4			
ID	AA013525		
AC	AA013525	standard: Protein: 221 AA.	
AC	AA013525:		
DE	06-NOV-2001 (first entry)		
DE	Human polypeptide SEQ ID NO 27417.		
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukemia;		
KW	nervous system disorders; arthritis; inflammation.		
OS	Homo sapiens.		
PN	WO200164835-A2.		
PN	07-SEP-2001.		
PD	26-FEB-2001; 2001WO-US04927.		
PF	28-FEB-2000; 2000US-0515126.		
PR	18-MAY-2000; 2000US-0577409.		
PR	(HYSE-) HXSEQ INC.		
PA	Tang YT, Liu C, Drmanac RF;		
PI	WPI: 2001-514838/56.		
PI	N-PSDB; AA193456.		
PT	Isolated nucleic acids and polypeptides, useful for preventing		
PT	diagnosing and treating e.g. leukemia, inflammation and immune		
PT	disorders -		
XX	Claim 20; SEQ ID NO 27417; 1399pp + Sequence Listing: English.		
PS	The invention relates to human polynucleotides (AA179941-AA193841) and		
CC	the encoded proteins (AA000010-AA013510) that exhibit activity elating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	e.g. stem cell growth factor activity, haematopoiesis regulating		
CC	activity, tissue growth factor activity, immunomodulatory activity and		
CC	activin/Inhbn activity and may be useful in the diagnosis and/or		
CC	treatment of cancer, leukemia, nervous system disorders, arthritis and		
CC	Inflammation.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WFO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 221 AA:		
XX	221 AA:		
Query Match	86.7%;	Score 768.5;	DB 22;
Best Local Similarity	76.5%;	Pred. No. 1,5e-69;	Length 221;
Matches 156;	Conservative 8;	Mismatches 7;	Indels 33;
		Gaps 2	
2	SSNNPYPYDLTKLLIGDGSVGRSCLLARFADYVYESYISTGCVFKIRTEIDGKTIK	61	
19	AAANPPYDTEKLLIGDGSVGRSCLLARFADYVYESYISTGCVFKIRTEIDGKTIK	78	





CC	infections, such as aplastic anaemia, ischaemic injuries such as myocardial
CC	infarction, stroke, and reperfusion injury, toxin-induced diseases such
CC	as alcohol-induced liver damage, cirrhosis, and lathyrism wasting
CC	diseases such as cachexia, viral infections, and osteoporosis. They can
CC	also be used to stimulate cell proliferation for use in transplantation
CC	or to produce cells to fight an infection or a cancer, or to correct a
CC	genetic defect in a disease such as sickle cell beta thalassaemia, cystic
CC	fibrosis or Huntington's chorea. Antagonists can be used to prevent or
CC	treat a disorder associated with cell proliferation e.g. cancers or
CC	inflammation, e.g. Addison's disease, adult respiratory distress
CC	syndrome, allergies, asthma, atherosclerosis, bronchitis, cholecystitis,
CC	Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis,
CC	diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis,
CC	gout, Graves' disease, hyperostosis/hypophosphatosis, irritable bowel syndrome, lupus
CC	erythematosus, multiple sclerosis, myasthenia gravis, inflammation,
CC	osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid
CC	arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis,
CC	complications of cancer, haemodialysis, extracorporeal circulation,
CC	infections and trauma.
CC	
XX	
SQ	Sequence 201 AA:
Query Match	86.5%; Score 766.5; DB 20; Length 201;
Best Local Similarity	77.2%; Pred. No. 2.1e-69;
Matches 156; Conservative	6; Mismatches 7; Indels 33; Gaps 2
OY	4 MNPEYVLEKLLIGDSCGSKSLLRADDTYESISTGCVFKRTIELDKTKTIO 63
DB	1 MNPEYVLEKLLIGDSCGSKSLLRADDTYESISTGCVFKRTIELDKTKTIO 60
OY	64 I-----ESNNYKQWLEIDRVSNNWKLNG 91
DB	61 IWDPAQGEFRFTISVYRCAGIILVYDWDTESYANQKMDRIDRASENNKLNG 120
OY	92 NKQDITTKRVVDYTTAKEPDSIGTPEFTSARNANVQSPFMAETKRNKGATAG 151
DB	121 NKSDLTTRKYVDNTTAEPFDSIGTPEFTSARNANVQSPFMAETKRNKGATAG 180
OY	152 GAESNVKTIQSPFKKQSGGGCC 173
DB	181 G-ENFMKIDTSPVKAPGGCC 201
RESULT 7	
AA028024	AA028024 standard; Protein: 201 AA.
AC	AA028024:
AC	AA028024:
DT	18-DEC-2001 (first entry)
XX	
DE	Novel human secretory protein, Seq ID No 193.
XX	
KM	Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KM	ischaemic reperfusion injury; haematopoiesis; cancer; neuropathy;
KM	transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KM	amylotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KM	ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KM	gut protection; lung; liver fibrosis; immune deficiency; infection;
KM	severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KM	multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KM	fertility; analgesic; pain; antigen.
OS	
XX	Homo sapiens.
PN	MO20016689-A2.
PN	MO20016689-A2.
PD	13-SEP-2001.
XX	
XX	05-MAR-2001; 2001WO-US04942.
XX	
XX	07-MAR-2000; 2000US-0519705.
RR	19-MAY-2000; 2000US-0574454.

PR	17-JUN-2000;	2000US-0956193.
PR	14-JUL-2000;	2000US-0616847.
PR	19-SEP-2000;	2000US-0663363.
PR	20-OCT-2000;	2000US-0693267.
XX		
FR	(HISE-) HXSD INC.	
FR		
XX	Tang YF, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;	
PL	Zhao QM, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;	
DR	WPI: 2001-589934/66.	
XX	N-PSTDB: AAS44924.	
PT	Novel polypeptides and nucleic acids obtained from cDNA libraries	
PT	prepared from various human tissues, for diagnosis and treatment of	
PT	cancer, neurological, inflammatory, and autoimmune disorders -	
XX		
PS	Example 3; SEQ ID NO 193; 107bp; English.	
XX		
CC	The invention relates to novel isolated human secreted polypeptides (I)	
CC	and polynucleotides (II). (I) and (II) are useful for treating	
CC	Inflammatory conditions such as arthritis, neuritis, Crohn's disease,	
CC	Ischaemia-reperfusion injury, shock, sepsis, immune responses, and is	
CC	involved in increasing haematopoiesis, stem cell survival, bone growth	
CC	and remodeling. (I), (II) and modulators of (II) are useful for	
CC	prophylaxis or treatment of one or more cancers. (II) is also useful for	
CC	creating transgenic animals useful for studying the in vivo activities of	
CC	the polypeptide as well as for studying modulators of the polypeptides.	
CC	(I) induces the proliferation of neural cells and regeneration of nerve	
CC	tissue and brain tissue and is useful for the treatment of central and	
CC	peripheral nervous system diseases and neuropathies, such as Alzheimer's,	
CC	Parkinson's disease, Huntington's disease, and amyotrophic lateral	
CC	sclerosis. In addition, (I) is involved in chemotactic or chemokinetic	
CC	activity, regulation of hematopoietists and is useful for treating myeloid	
CC	or lymphoid cell disorders, platelet disorders such as thrombocytopenia	
CC	and for regeneration of bone, cartilage, tendon, ligament and/or nerve	
CC	tissue growth, and in tissue repair, healing of burns, incisions,	
CC	ulcers, for treating osteoporosis, osteoarthritis, bone degenerative	
CC	disorders, or periodontal disease. Furthermore, (I) is also useful for	
CC	gut protection or regeneration and treatment of lung or liver fibrosis,	
CC	reperfusion injury in various tissues, various immune deficiencies and	
CC	disorders including severe combined immunodeficiency (SCID), bacterial or	
CC	fungal infections, autoimmune disorders e.g. multiple sclerosis,	
CC	rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic	
CC	reactions and conditions, such as asthma or other respiratory problems.	
CC	In addition, (I) affects biorhythms or circadian cycles of rhythms,	
CC	of fertility, metabolism, catabolism, anabolism, storage or elimination of	
CC	dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides	
CC	analgesic effects or other pain reducing effects, immunoglobulin like	
CC	activity and can act as an antigen in a vaccine composition to raise an	
CC	immune response. AAU28020-AAU28395 represent novel human secreted protein	
XX	amino acid sequences of the invention.	
XX		
SQ	Sequence 201 AA:	
	Query Match 86.5% Score 766.5; DB 22; Length 201:	
	Best Local Similarity 77.2%; Pred. No. 2.le-69;	
	Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2	
OY	4 MNPEVDYFKLLIGDSGVKSKCLLRADPDYTESYSTIGVDFKIRITIEDKTIKLQ 63	
DB	1 MNPEVDYFKLLIGDSGVKSKCLLRADPDYTESYSTIGVDFKIRITIEDKTIKLQ 60	
OY	64 I-----ESFNVMKOMLOEIDRYASENNKVLVG 91	
DB	61 IMDTAGQENRFTITSYRGANGIIVYDVTDDSEYANVMKOMLOEIDRYASENNKVLVG 120	
OY	92 NKCDLTKRKVVYDTAAEPADSILGIPLETSAKNATNEOQSFTNMAAEIKRMKPATGA 151	
DB	121 NKSDDLTKRKVVNDTAAEPADSILGIPLETSAKNATNEOQAFTNMAAERIKRMKPGCA 180	
OY	152 GAERSNVKTOSTPVKGSGGCC 173	

Db 181 G-ERRNLIKIDSTPVKPGAGGCC 201

RESULT 8  
ID ABB59808 standard; Protein: 205 AA.  
XX  
AC ABB59808;  
XX  
D7 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 6216.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX DR N-PSDB; ABL03911.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 6216; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (AB57737-AB572072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 205 AA;

Query Match 77.1%; Score 683; DB 22; Length 205;  
Best Local Similarity 69.4%; Fred. No. 6e-61;  
Matches 143; Conservative 12; Mismatches 17; Indels 34; Gaps 3;

QY 1 MSNMPEYDLFKLLIGDSGVGKSCLLFRADPTYSYISIGVDKRIETLDGKTI 60  
Db 1 MSSVNPEDYLFKLLIGDSGVGKSCLLFRADPTYSYISIGVDKRIETLDGKTI 60  
QY 61 KLGIT-----ESFNWYKWLQEDIRYASENVKL 88  
Db 61 KLGITDAGGERFRTISSTYRGAGHILVYDCTDQSEFNWYKWLQEDIRYASENVKL 120  
QY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNAITNVEOSPTMAAEIKRMGPGA 148  
Db 121 LVGNKSDLTTRKVDHTTAETAAOLGIPLETSAKNAITNVEOAFMTAAEIKRMVPPS 180  
QY 149 TAGGAESNWKI-OSTPVKOSGGGCC 173  
Db 181 SATD-NASKVKIIDGRPVNTKSGCC 205

RESULT 9  
ID AAG07763  
XX  
AC AAG07763;  
XX  
D7 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 5040.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 07-MAY-1999; 99US-0132487.  
XX 11-MAY-1999; 99US-0132863.  
XX 14-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134376.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136783.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139493.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.

```
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140623.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147419.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158282.
PR 12-OCT-1999; 99US-0158360.
PR 13-OCT-1999; 99US-0158363.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159285.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 71 94; Score 637; DB 21; Length 202;
Beat Local Similarity 65 34; Pred No. 2 6e-56;
Matches 133; Conservative 11; Mismatches 25; Indels 34; Gaps 3;

QY 4 MNEPDYLFKLLIGSGGKSCILLRADDYTBESYISTYICVDFKRTIELDKTILQ 63
DB 1 MNEPDYLFKLLIGSGGKSCILLRADDYTBESYISTYICVDFKRTIELDKTILQ 60
QY 64 I-----ESFNNTKQMLQEDIRVASENVKLLVG 91
DB 61 IMDYGAQREPTITSSYKANGIIVYDVTDLSESNNTKQMLQEDIRVASENVKLLVG 120
QY 92 NKCDLTKKVVDTTAKFADSLGIPLETSAARNATVDSQFMTMAELIKRKGCGATAG 151
```

```

Db      121 NKCGLTSQKRVSTETAKAFADDELGIPELETSAKNATNWEALPMANTPAIKTRMA-SQAPG 179
OY      152 GAENSNVRIQSTPVKOSGGCC 173
        :::::|::|::|
Db      180 GSKPPTVQIRGQPVNQ-SGCC 200

RESULT 10
AAG10858
ID      AAG10858 standard; Protein; 202 AA.
XX
AC      AAG10858;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 9348.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KM      hybridisation assay; genetic mapping; gene expression control; promoter;
XX      termination sequence.
OS      Arabidopsis thaliana.
XX
PN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PE      25-FEB-2000; 2000EP-0301439.
XX
PR      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      23-APR-1999; 99US-0130891.
PR      28-APR-1999; 99US-0131449.
PR      30-APR-1999; 99US-0132048.
PR      04-MAY-1999; 99US-0132407.
PR      05-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.

PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142927.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      11-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.

```

PR 13-AUG-1999: 99US-0148565.  
PR 13-AUG-1999: 99US-0148684.  
PR 16-AUG-1999: 99US-0149368.  
PR 17-AUG-1999: 99US-0149375.  
PR 18-AUG-1999: 99US-0149426.  
PR 20-AUG-1999: 99US-0149722.  
PR 20-AUG-1999: 99US-0149723.  
PR 20-AUG-1999: 99US-0149929.  
PR 23-AUG-1999: 99US-0149902.  
PR 23-AUG-1999: 99US-0149930.  
PR 25-AUG-1999: 99US-0150566.  
PR 26-AUG-1999: 99US-0150884.  
PR 27-AUG-1999: 99US-0151065.  
PR 27-AUG-1999: 99US-0151066.  
PR 30-AUG-1999: 99US-0151080.  
PR 31-AUG-1999: 99US-0151303.  
PR 01-SEP-1999: 99US-0151330.  
PR 07-SEP-1999: 99US-0152363.  
PR 10-SEP-1999: 99US-0153070.  
PR 13-SEP-1999: 99US-0153758.  
PR 15-SEP-1999: 99US-0154039.  
PR 20-SEP-1999: 99US-0154779.  
PR 22-SEP-1999: 99US-0155139.  
PR 23-SEP-1999: 99US-0155486.  
PR 24-SEP-1999: 99US-0155659.  
PR 28-SEP-1999: 99US-0156458.  
PR 29-SEP-1999: 99US-0156596.  
PR 04-OCT-1999: 99US-0157117.  
PR 05-OCT-1999: 99US-0157753.  
PR 06-OCT-1999: 99US-0157865.  
PR 07-OCT-1999: 99US-0158029.  
PR 08-OCT-1999: 99US-0158232.  
PR 12-OCT-1999: 99US-0158369.  
PR 13-OCT-1999: 99US-0159283.  
PR 13-OCT-1999: 99US-0159293.  
PR 13-OCT-1999: 99US-0159329.  
PR 14-OCT-1999: 99US-0159330.  
PR 14-OCT-1999: 99US-0159331.  
PR 14-OCT-1999: 99US-0159637.  
PR 14-OCT-1999: 99US-0159638.  
PR 18-OCT-1999: 99US-0159584.  
PR 21-OCT-1999: 99US-0160741.  
PR 21-OCT-1999: 99US-0160767.  
PR 21-OCT-1999: 99US-0160768.  
PR 21-OCT-1999: 99US-0160770.  
PR 21-OCT-1999: 99US-0160814.  
PR 21-OCT-1999: 99US-0160815.  
PR 22-OCT-1999: 99US-0160980.  
PR 22-OCT-1999: 99US-0160981.  
PR 22-OCT-1999: 99US-0161404.  
PR 25-OCT-1999: 99US-0161405.  
PR 25-OCT-1999: 99US-0161406.  
PR 25-OCT-1999: 99US-0161359.  
PR 26-OCT-1999: 99US-0161360.  
PR 26-OCT-1999: 99US-0161361.  
PR 28-OCT-1999: 99US-0161920.  
PR 28-OCT-1999: 99US-0161992.  
PR 28-OCT-1999: 99US-0161993.  
PR 29-OCT-1999: 99US-0162142.

Query Match 71.4% Score 633: DB 21: Length 202;  
Best Local Similarity 65.8% Pred. No 6.7e-56;  
Matches 133; Conservative 10; Mismatches 25; Indels 34; Gaps 3;  
DB 4 NRPEDYFKLLIGSGVSKSLLPADDTYESYISTGVDFKRTIEDGKTIKQ 63  
DB 1 NRPEDYFKLLIGSGVSKSLLPADDTYESYISTGVDFKRTIEDGKTIKQ 60  
QY 64 I-----ESFNNAKQMLQEDIRYASENKKLLVG 91

DB 61 IWDPAQGEFRTITSSYRCAGIIVTVDLESPNNNAQWLNEDIDTASENKKLLVG 120  
QY 92 NKCDLTKKVVDDYTTAKEFPADSLGIPLETSAKKAINVEDQSPMTAAEIKRGPATAG 151  
DB 121 NKNDLTSQKRVSTETAAAFADGLGIPLETSAKKATVEEAPMAATTAIKTRAA-SQPA 179  
QY 152 GAERSNVKIQSTPVKQSGGGCC 173  
DB 180 GAKPFTVQIRGQPVNQO-SGCC 200  
RESULT 11  
AAG30498  
ID AAG30498 standard; Protein: 258 AA.  
AC AAG30498;  
DT 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36472.  
XX  
XX Protein identification; signal transduction pathway: metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
XX EPI033405-A2.  
PN  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PR 25-FEB-1999: 99US-0121825.  
PR 05-MAR-1999: 99US-0123160.  
PR 09-MAR-1999: 99US-0123548.  
PR 23-MAR-1999: 99US-0125788.  
PR 25-MAR-1999: 99US-0126264.  
PR 29-MAR-1999: 99US-0126785.  
PR 01-APR-1999: 99US-0127462.  
PR 06-APR-1999: 99US-0128234.  
PR 08-APR-1999: 99US-0128714.  
PR 16-APR-1999: 99US-0129845.  
PR 19-APR-1999: 99US-0130077.  
PR 21-APR-1999: 99US-0130449.  
PR 23-APR-1999: 99US-0130510.  
PR 28-APR-1999: 99US-0130891.  
PR 30-APR-1999: 99US-0131449.  
PR 30-APR-1999: 99US-0132048.  
PR 04-MAY-1999: 99US-0132407.  
PR 05-MAY-1999: 99US-0132484.  
PR 06-MAY-1999: 99US-0132485.  
PR 07-MAY-1999: 99US-0132486.  
PR 11-MAY-1999: 99US-0132863.  
PR 14-MAY-1999: 99US-0134256.  
PR 14-MAY-1999: 99US-0134218.  
PR 14-MAY-1999: 99US-0134221.  
PR 14-MAY-1999: 99US-0134370.  
PR 18-MAY-1999: 99US-0134768.  
PR 19-MAY-1999: 99US-0134941.  
PR 20-MAY-1999: 99US-0135124.  
PR 21-MAY-1999: 99US-0135353.  
PR 24-MAY-1999: 99US-0135829.  
PR 25-MAY-1999: 99US-0136021.  
PR 27-MAY-1999: 99US-0136582.  
PR 28-MAY-1999: 99US-0136752.  
PR 01-JUN-1999: 99US-0137222.  
PR 03-JUN-1999: 99US-0137528.  
PR 04-JUN-1999: 99US-0137502.  
PR 07-JUN-1999: 99US-0137724.

PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 17-JUN-1999; 99US-0139453.  
 PR 18-JUN-1999; 99US-0139452.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 21-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144886.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148365.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149175.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0149930.  
 PR 26-AUG-1999; 99US-0150864.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 69.9%; Score 619.5; DB 21; Length 258;  
 Best Local Similarity 61.6%; Pred. No. 2,le-54;  
 Matches 125; Conservative 15; Mismatches 30; Indels 33; Gaps 2;

```
OY 3 SNNPEYDLEFKLLIDSGVSKSCLLRFADDTYESITITGVDEKRTIIELODKTL 62
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 55 TNNPEYDLEFKLLIDSGVSKSCLLRFSDDBYESTISTIGVDEKRTIIELODKTL 114
OY 63 OI-----EFNNWQWIOEIDRASENNKLLV 90
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 115 QIMWDAQGERRTITSSYRCAHGIIIVYDVDESEFNWKMISEIDRTASNNKLLV 174
OY 91 GAKCULTTKKVVDTTAKAFADSLGIPLETSSKATNVEQSEFTMALETIKRKGATGA 150
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 175 GKNSDLTNNRFLPRTAKAFADSLGIPLETSSKATNVEQSEFTMALETIKRKGATGA 234
OY 151 GGAENSNKXIOSTPVKQSGGCC 173
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 235 NNAAPPYQIRGQPVAK-NGCC 256
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 12
AAG30499
ID AAG30499 standard; Protein; 203 AA.
XX
AC AAG30499;
XX
Dt 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36473.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
PD 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121325.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123948.
PR 22-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130849.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135529.
PR 25-MAY-1999; 99US-0135621.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140658.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140861.
PR 01-JUL-1999; 99US-0141267.
PR 02-JUL-1999; 99US-0141842.
PR 06-JUL-1999; 99US-0142053.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
```

PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147760
PR	06-AUG-1999	9905-01474103
PR	06-AUG-1999	9905-0147763
PR	09-AUG-1999	9905-0147933
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148319
PR	12-AUG-1999	9905-0148841
PR	13-AUG-1999	9905-0148565
PR	16-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149722
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	23-AUG-1999	9905-0150684
PR	26-AUG-1999	9905-0150884
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151338
PR	01-SEP-1999	9905-0151263
PR	07-SEP-1999	9905-0152363
PR	10-SEP-1999	9905-0153070
PR	13-SEP-1999	9905-0153758
PR	15-SEP-1999	9905-0154019
PR	16-SEP-1999	9905-0154039
PR	20-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155133
PR	22-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0156559
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157573
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158362
PR	13-OCT-1999	9905-0158293
PR	13-OCT-1999	9905-0159294
PR	14-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159331
PR	14-OCT-1999	9905-0159637
PR	18-OCT-1999	9905-0159584
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160778
PR	21-OCT-1999	9905-0160981
PR	22-OCT-1999	9905-0161004
PR	25-OCT-1999	9905-0161404
PR	25-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	26-OCT-1999	9905-0161356
PR	26-OCT-1999	9905-0161360
PR	26-OCT-1999	9905-0161361
PR	28-OCT-1999	9905-0161992
PR	28-OCT-1999	9905-0161992
PR	29-OCT-1999	9905-0162142

Query Match	Similarity	69.8%	Score 618.5	DB 21	Length 203
Best Local	Similarity	61.9%	Pred. No. 2e-54		
Matches	125	Conservative	14	Mismatches	30
				Indels	33
				Gaps	2
QY	4	MMPEVDYFLLKLLIDSGVKSCLLRFADADVTYESYSTIGVDFFKRTIELDGKTIKQ	63		
DB	1	MMPEVDYFLLKLLIDSGVKSCLLRFADADVTYESYSTIGVDFFKRTIELDGKTIKQ	60		
QY	64	I-----ESFNMYKQMLQETDRYASENVKLLVG	91		
DB	61	IWDTAGOERFRTTSSYYRGAGHGIITVYDVTDESEFNMYKQMLSEIDRYASNVNKLWVG	120		
QY	92	NKCDLTKKVVYDTTAKERFADSLGIPELETSAKNATNVQSPMTMAAEIKKRMGPGATAG	151		
DB	121	NKSDLTENNAIPEYETAKAFADDELGIPEMETSAKDNTNVEQATMAASATIKENASOPAGN	180		
QY	152	GAEKSNVRIQSTPVKQSGGCC	173		
DB	181	NARPTVQIRGQPVAK-NGCC	201		
RESULT	13				
ID	AAG30710				
XX	AAG30710 standard; Protein; 218 AA.				
XX	AAG30710;				
XX					
DT	17-OCT-2000 (first entry)				
XX					
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 36764.				
XX					
KM	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
XX	termination sequence.				
OS	Arabidopsis thaliana.				
XX					
PN	EP1033405-A2.				
XX					
PD	06-SEP-2000.				
XX					
PF	25-FEB-2000; 2000EP-0301439.				
XX					
PR	25-FEB-1999; 990S-0121825.				
PR	05-MAR-1999; 990S-0123180.				
PR	09-MAR-1999; 990S-0122548.				
PR	23-MAR-1999; 990S-0125788.				
PR	25-MAR-1999; 990S-0126264.				
PR	29-MAR-1999; 990S-0126785.				
PR	01-APR-1999; 990S-0127462.				
PR	06-APR-1999; 990S-0128234.				
PR	08-APR-1999; 990S-0128714.				
PR	16-APR-1999; 990S-0129845.				
PR	19-APR-1999; 990S-0130077.				
PR	21-APR-1999; 990S-0130449.				
PR	23-APR-1999; 990S-0130510.				
PR	23-APR-1999; 990S-0130891.				
PR	28-APR-1999; 990S-0131449.				
PR	30-APR-1999; 990S-0132048.				
PR	30-APR-1999; 990S-0132407.				
PR	04-MAY-1999; 990S-0132484.				
PR	05-MAY-1999; 990S-0132485.				
PR	06-MAY-1999; 990S-0132486.				
PR	06-MAY-1999; 990S-0132487.				
PR	07-MAY-1999; 990S-0132863.				
PR	11-MAY-1999; 990S-0134256.				
PR	14-MAY-1999; 990S-0134218.				
PR	14-MAY-1999; 990S-0134219.				
PR	14-MAY-1999; 990S-0134221.				
PR	14-MAY-1999; 990S-0134370.				
PR	18-MAY-1999; 990S-0134768.				
PR	19-MAY-1999; 990S-0134941.				



PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135829.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136982.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139850.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142290.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 23-JUL-1999; 99US-0145082.  
PR 23-JUL-1999; 99US-0145082.  
PR 23-JUL-1999; 99US-0145082.  
PR 23-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 26-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149272.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0150303.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 07-SEP-1999; 99US-0152363.  
PR 07-SEP-1999; 99US-0153707.  
PR 10-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159293.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159337.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 21-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
Query Match 69.1%; Score 612; DB 21; Length 218;  
Best Local Similarity 65.0%; Pred. No. 9,8e-54;  
Matches 128; Conservative 11; Mismatches 24; Indels 34; Gaps 3;  
OY 9 DYLFKLLIGDSGVKSCLLRFADDTTESTISTIGDFKRTTELGGTKIKLQI----- 64  
DB 22 DYLFKLLIGDSGVKSCLLRFADDTSTIDSTISTIGDFKRTTELGGTKIKLQIMPTA 81  
OY 65 -----ESFNNVKNQMLQELIDRYASENNKLLVGNKCDL 96  
DB 82 GOERFRITSSYRGAAGLIYTVDTLDESFNNNVKNQMLNEIDRYASENNKLLVGNKCDL 141  
OY 97 TTKKVVDTTAKEPADSLGIPLETSAKNATNVEOSEFTMAAEIKKRGPGATGAGAEKS 156  
DB 142 TSOKVSTETAKAFADDELGIPELETSAKNATNVEAFAMATKRTMA-SQPAAGSKRP 200  
OY 157 NKKIOSTPVKSGGSC 173  
DB 201 TVQIRGQPVNQ-SGCC 216  
RESULT 14  
AAG30709  
ID AAG30709 standard; Protein; 221 AA.  
AC AAG30709;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36763.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
PD  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140333.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.

```
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 15-AUG-1999; 99US-0148684.
PR 15-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 23-OCT-1999; 99US-0160980.
PR 23-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.1%; Score 612; DB 21; Length 221;
Best Local Similarity 65.0%; Pred No. 1e-53;
Matches 128; Conservative 11; Mismatches 24; Indels 34; Gaps 3;

QY 9 DYEKLLIGDSGVKSCLLRPADDTYESYISTGVDFKRTIEHDKTIRLOI---- 64
DB 25 DYEKLLIGDSGVKSCLLRPADDTYESYISTGVDFKRTIEHDKTIRLOI 84
QY 65 -----ESFNKKQMLQEDIRASENVKLLVGNKCDL 96
DB 85 GDRFRTTSSYRGAGIIVTYDVTLESEFNKKQMLQEDIRASENVKLLVGNKCDL 144
QY 97 TTKKVVDTYTAKEPADSLGIPELETSAAKNATNVEQSPTMAAEIKKRMKGATAGGAERK 156
DB 145 TSOKVSTETAKAPADELGIPLETSAAKNATNVEAFAMMTAIRTMA-SQAPAGSRFP 203
QY 157 NKRIGSTPYKQSGGCC 173
DB 204 TVOIRGQPVNOQ-SGCC 219

RESULT 15
AAC45323
AAC45323 standard; Protein: 254 AA.
ID AAC45323;
AC AAC45323;
XX 18-OCT-2000 (first entry).
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56887.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
```

PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140699.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143622.
PR	15-JUL-1999;	99US-0144003.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145087.
<hr/>		
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150888.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151428.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	13-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.

PR 21-OCT-1999; 990S-0160814.  
 PR 21-OCT-1999; 990S-0160815.  
 PR 22-OCT-1999; 990S-0160980.  
 PR 22-OCT-1999; 990S-0160981.  
 PR 22-OCT-1999; 990S-0160989.  
 PR 25-OCT-1999; 990S-0161404.  
 PR 25-OCT-1999; 990S-0161405.  
 PR 25-OCT-1999; 990S-0161406.  
 PR 26-OCT-1999; 990S-0161359.  
 PR 26-OCT-1999; 990S-0161360.  
 PR 26-OCT-1999; 990S-0161361.  
 PR 28-OCT-1999; 990S-0161920.  
 PR 28-OCT-1999; 990S-0161992.  
 PR 28-OCT-1999; 990S-0161993.  
 PR 29-OCT-1999; 990S-0162142.

Query Match 69.18; Score 612; DB 21; Length 254;  
 Best Local Similarity 65.08; Pred No. 1,2e-53;  
 Matches 128; Conservative 11; Mismatches 24; Indels 34; Gaps 3;

OY 9 DYIFKLLIGDSGVGKSCILLRPAADTYESYTGVDKFKRTIELDKTIKQI---- 64  
 Db 58 DYIFKLLIGDSGVGKSCILLRPAADTYESYTGVDKFKRTIELDKTIKQIIMDRA 117  
 OY 65 -----ESPNNYKQHLQIDRYASENKKLVGNKCDL 96  
 Db 118 GQERFRITTSYRGAGHIIYTYDVTLESPNNYKQHLNEIDRYASENKKLVGNKCDL 177  
 OY 97 TTKKVVDTYTAKEPADSLGIPLETSAKNATUNVQSFMTMAELIKRKGPGATAGGAERK 156  
 Db 178 TSKRVSTETAKAFADSLGIPLETSAKNATUNVQSFMTMAELIKRKGPGATAGGAERK 236  
 OY 157 NVKIGSTPVKQSGGGCC 173  
 Db 237 TVQIRGQPVNOQ-SGCC 232

Search completed: June 18, 2003, 15:21:09  
 Job time : 72 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using SW model

Run on: June 18, 2003, 15:19:09 ; Search time 41 Seconds  
(without alignments)  
405,640 Million cell updates/sec

Title: US-09-820-003b-2

Perfect score: 886  
Sequence: 1 MSSNPEYDYLFKLLIGDS.....EKSNNKIQSTPVKSGGCC 173

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	860	97.1	205	1	TVBQYP
2	860	97.1	205	1	TVBQYP
3	860	97.1	205	1	TVBQYP
4	849	95.8	205	1	TVBQYP
5	812.5	91.7	201	2	D38625
6	771.5	87.1	201	2	S06147
7	733	82.7	205	2	S38339
8	671	75.7	202	2	JR0318
9	660	74.5	205	2	T33781
10	658.5	74.3	203	2	S30096
11	646.5	73.0	203	2	JC4105
12	646.5	73.0	203	2	T50323
13	642.5	72.5	203	2	JC3247
14	631.5	71.5	201	2	UC3237
15	627.5	70.8	203	2	S34253
16	623.5	70.4	206	2	S04599
17	619.5	69.9	226	2	B86133
18	619.5	69.9	202	2	S41430
19	618.5	69.8	203	2	B38202
20	613.5	69.2	202	2	S38740
21	612	69.1	221	2	H71444
22	598	67.5	218	2	T07609
23	590.5	66.6	202	2	S72515
24	587	66.3	201	2	S39565
25	574	64.8	206	2	T14391
26	572.5	64.6	196	2	PS0279
27	571	64.4	206	1	TVBYQ2
28	523	59.0	208	2	A38202
29	468.5	53.9	208	2	A34716

30	461	52.0	203	2	B34716	GRP-binding protei
31	447	50.5	216	2	T48378	GRP-binding protei
32	442	49.9	215	2	T14565	GRP-binding protei
33	439.5	49.6	215	2	S57478	GRP-binding protei
34	438.5	49.5	216	2	J50640	GRP-binding protei
35	435.5	49.2	200	2	S12790	GRP-binding protei
36	430.5	48.6	215	2	S57462	GRP-binding protei
37	429.5	48.5	216	2	S57471	GRP-binding protei
38	429.5	48.5	222	2	T14405	small GRP-binding
39	428.5	48.4	216	2	T45901	GRP-binding protei
40	426.5	48.1	215	2	S57474	Grpase ALRAB8 - Ar
41	425.5	48.0	216	2	S33900	GRP-binding protei
42	418	47.2	207	2	B49647	GRP-binding protei
43	418	47.2	207	2	B36364	GRP-binding protei
44	410.5	46.3	206	2	T78851	GRP-binding protei
45	409.5	46.2	209	2	B38625	GRP-binding protei

## ALIGNMENTS

## RESULT 1

TVBQYP GRP-binding protein Rab1 - human

N:Alternate names: protein DKFZP564B163.1

C:Species: Homo sapiens (man)

C>Date: 29-Jun-1990 #sequence\_rev1sion 06-Dec-1996 #text\_change 19-Jan-2001

C:Accession: A34323; T08698

R:Zahrouk, A.; Touchot, N.; Chardin, P.; Tavittan, A.

J. Biol. Chem. 264, 12394-12401, 1989

A:Title: The human Rab genes encode a family of GTP-binding proteins related to yeast

A:Reference number: A34323; MUID:89308668; PMID:2501306

A:Accession: A34323

A:Molecule type: mRNA

A:Residues: 1-205 <ZMH>

A:Cross-references: GB:J04941; GB:M28209; NID:9550059; PIDN:AAA60240.1; PID:9550060

A:Experimental source: phaeochromocytoma

R:Wambitt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z16471

A:Accession: T08698

A:Molecule type: mRNA

A:Residues: 1-64,141-205 <MAN>

A:Cross-references: EMBL:AL050268

A:Experimental source: fetal brain; clone DKFZP564B163

C:Genetics:

A:Gene: GDB:RAB1

A:Cross-references: GDB:118857; OMIM:179508

A:Map position: Ap15.31-4p15.31

A:Note: DKFZP564B163.1

C:Function:

A:Description: probably involved in protein transport from the endoplasmic reticulum

A:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; Lipoprotein; membran

tein

F:1-205/Product: GRP-binding protein Rab1 #status predicted <MAN>  
F:1-64,141-205/Product: GRP-binding protein Rab1, splice variant #status predicted <M  
F:12-117/Domain: translation elongation factor Tu homology <ETU>  
F:18-23/Region: nucleotide-binding motif A (P-loop)  
F:62-67/Region: nucleotide-binding motif B  
F:124-127/Region: GTP-binding NKXD motif  
F:154-156/Region: GTP-binding SAK/L motif  
F:24,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #  
F:194/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status predicted  
F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status experimental

Query Match 97.1% Score 860; DB 1; Length 205;

Best local similarity 84.4% Pred. No. 2,2e-62;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

DB 1 MSSNPEYDYLFKLLIGDSGVGSCILLRFADYTESYTGVDFTIRIELDKRTI 60  
1 MSSNPEYDYLFKLLIGDSGVGSCILLRFADYTESYTGVDFTIRIELDKRTI 60

```

QY 61 KIQI-----ESFNNVKQWLOEIDRYASENVKL 88
    |||
Db 61 KQIWDTAGQERFRTITSSYRGAGIIVYDVDTQESFNNVKQWLOEIDRYASENVKL 120
QY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKKRMGPGA 148
Db 121 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKKRMGPGA 180
QY 149 TAGGAEKSNVKIQSTPVKQSGGCC 173
Db 181 TAGGAEKSNVKIQSTPVKQSGGCC 205

```

## RESULT 2

```

TYDGYF
GTP-binding protein Rab1 - dog
C:Species: Canis lupus familiaris (dog)
C>Date: 19-Feb-1994 #sequence_revision 06-Dec-1996 #text_change 19-Jan-2001
C:Accession: S19104; A36364; S15600
R:Zerial, M.
submitted to the EMBL Data Library, August 1990
A:Reference number: S19104
A:Accession: S19104
A:Molecule type: mRNA
A:Residues: 1-205 <ZER>
A:Cross-references: EMBL:X56384
R:Chavrier, P.; Vignon, M.; Sander, C.; Simons, K.; Zerial, M.
Mol. Cell. Biol. 10, 6578-6585, 1990
A:Title: Molecular cloning of YP1/SEC4-related CDNs from an epithelial cell line.
A:Reference number: A36364; MUID:91061765; PMID:2123294
A:Accession: A36364
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-157; 'EK', 160-205 <CHA>
A:Cross-references: GB:X56384; NID:9913
C:Function:
A:Description: probably involved in protein transport from the endoplasmic reticulum thr
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipoprotein; membrane f
tein
F:12-127/Domain: translation elongation factor Tu homology <ERT>
F:18-25/Region: nucleotide-binding motif A (P-loop)
F:62-67/Region: nucleotide-binding motif B
F:124-127/Region: GTP-binding NKX motif
F:154-156/Region: GTP-binding SAK/L motif
F:124,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
F:194/Binding site: phosphate (Ser) (covalent) (by cd2 kinase) #status predicted
F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

```

Query Match 97.1%; Score 860; DB 1; Length 205;

Best Local Similarity 84.4%; Pred. No. 2.2e-62;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

```

QY 1 MSSNPEYDLFKILLIDSGVSKLLRFADDTYSYSTIGVDKIRITIEDGKI 60
Db 1 MSSNPEYDLFKILLIDSGVSKLLRFADDTYSYSTIGVDKIRITIEDGKI 60
QY 61 KIQI-----ESFNNVKQWLOEIDRYASENVKL 88
    |||
Db 61 KQIWDTAGQERFRTITSSYRGAGIIVYDVDTQESFNNVKQWLOEIDRYASENVKL 120
QY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKKRMGPGA 148
Db 121 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKKRMGPGA 180
QY 149 TAGGAEKSNVKIQSTPVKQSGGCC 173
Db 181 TAGGAEKSNVKIQSTPVKQSGGCC 205

```

## RESULT 3

TYMSYP

GTP-binding protein ypt1 - mouse

N:Alternate names: GTP-binding protein Rab1; ras-related protein ypt1; transforming p

C:Species: Mus musculus (house mouse)

C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 19-Jan-2001

C:Accession: S05551; S06285

R:Wichmann, H.; Disela, C.; Haubruck, H.; Gallwitz, D.

Nucleic Acids Res. 17, 6737-6738, 1989

A:Title: Nucleotide sequence of the mouse ypt1 gene encoding a ras-related GTP-bindin

A:Reference number: S05551; MUID:89386011; PMID:2506528

A:Accession: S05551

A:Molecule type: DNA

A:Residues: 1-205 <MIC>

A:Cross-references: EMBL:X15744; NID:955458; PIDN:CAA33760.1; PID:9763158

R:Haubruck, H.; Disela, C.; Wagner, P.; Gallwitz, D.

EMBO J. 6, 4049-4053, 1987

A:Title: The ras-related ypt protein is an ubiquitous eukaryotic protein: isolation a

A:Reference number: S06285; MUID:8816649; PMID:3127202

A:Accession: S06285

A:Molecule type: mRNA

A:Residues: 1-205 <HAU>

C:Genetics:

A:Gene: ypt1

A:Introns: 8/2; 32/3; 64/3; 96/3; 140/3

C:Function:

A:Description: probably involved in protein transport from the endoplasmic reticulum

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipoprotein; membran

tein

F:12-127/Domain: translation elongation factor Tu homology <ERT>

F:18-25/Region: nucleotide-binding motif A (P-loop)

F:62-67/Region: nucleotide-binding motif B

F:124-127/Region: GTP-binding NKX motif

F:154-156/Region: GTP-binding SAK/L motif

F:124,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #

F:194/Binding site: phosphate (Ser) (covalent) (by cd2 kinase) #status predicted

F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 97.1%; Score 860; DB 1; Length 205;

Best Local Similarity 84.4%; Pred. No. 2.2e-62;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

```

QY 1 MSSNPEYDLFKILLIDSGVSKLLRFADDTYSYSTIGVDKIRITIEDGKI 60
Db 1 MSSNPEYDLFKILLIDSGVSKLLRFADDTYSYSTIGVDKIRITIEDGKI 60
QY 61 KIQI-----ESFNNVKQWLOEIDRYASENVKL 88
    |||
Db 61 KQIWDTAGQERFRTITSSYRGAGIIVYDVDTQESFNNVKQWLOEIDRYASENVKL 120
QY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKKRMGPGA 148
Db 121 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKKRMGPGA 180
QY 149 TAGGAEKSNVKIQSTPVKQSGGCC 173
Db 181 TAGGAEKSNVKIQSTPVKQSGGCC 205

```

## RESULT 4

TYRTYP

GTP-binding protein Rab1 - rat

N:Alternate names: transforming protein ypt1 homolog

C:Species: Rattus norvegicus (Norway rat)

C>Date: 03-Aug-1992 #sequence\_revision 06-Dec-1996 #text\_change 19-Jan-2001

C:Accession: A39963

R:Touche, N.; Chardin, P.; Tavittian, A.

Proc. Natl. Acad. Sci. U.S.A. 84, 8210-8214, 1987

A:Title: Four additional members of the ras gene superfamily isolated by an oligonuci

A:Reference number: A39963; MUID:88068563; PMID:3317403

A:Accession: A39963

A:Molecule type: mRNA

A:Residues: 1-205 <TOU>

A:Cross-references: GB:J02998; NID:9206552; PIDN:AAA42006.1; PID:9206553



C:Function: A:Description: probably involved in protein transport from the endoplasmic reticulum th C:Superfamily: ras transforming protein; translation elongation factor Tu homology C:Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipoprotein; membrane F  
tein  
F:12-127/Domain: translation elongation factor Tu homology <ETU>  
F:18-25/Region: nucleotide-binding motif A (P-loop)  
F:62-67/Region: nucleotide-binding motif B  
F:124-127/Region: GTP-binding NKXD motif  
F:154-156/Region: GTP-binding SAK/L motif  
F:24,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asp, Lys, Asp, Ser) #sta  
F:194/Binding site: phosphate (Ser) (covalent) (by cdcc kinase) #status predicted  
F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 95.8%; Score 849; DB 1; Length 205;  
Best Local Similarity 83.4%; Pred. No. 1,7e-61;  
Matches 171; Conservative 0; Mismatches 2; Indels 32; Gaps 1;

OY 1 MSNPEYDYLFKLLIGDSGVKSCILLRFADDTYESYSTIGVDFKRTIELDGKTI 60  
DB 1 MSNPEYDYLFKLLIGDSGVKSCILLRFADDTYESYSTIGVDFKRTIELDGKTI 60  
OY 61 KIQI-----ESFNNVKOMLOEIDRYASENVNKL 88  
DB 61 KIQIDTAGQERFRRTTSYRGAGIIIVYDVDOESFNNVKOMLOEIDRYASENVNKL 120  
OY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRKGPGA 148  
DB 121 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRKGPGA 180  
OY 149 TAGGAESNVKIOSTPVKSGGCGC 173  
DB 181 TAGGAESNVKIOSTPVKSGGCGC 205

RESULT 5  
D38625  
GTP-binding protein o-rab1 - electric ray (Discopyge ommata)  
C:Species: Discopyge ommata  
C:Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 02-Feb-2001  
C:Accession: D38625  
R:Nggee, J.K.; Elferink, L.A.; Scheller, R.H.  
J. Biol. Chem. 266, 2675-2680, 1991  
A:Title: A family of ras-like GTP-binding proteins expressed in electrorotor neurons.  
A:Reference number: A38625; MUID:91115900; PMID:1899244  
A:Accession: D38625  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-201 <NGS>  
A:Cross-references: GB:M38393  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop:  
F:124/Domain: translation elongation factor Tu homology <ETU>  
F:15-22/Region: nucleotide-binding motif A (P-loop)  
F:121-124/Region: GTP-binding NKXD motif  
F:150-152/Region: GTP-binding SAK/L motif  
F:21,22,40,121,122,124,150/Binding site: Mg-GTP (Lys, Ser, Thr, Asp, Lys, Asp, Ser) #sta  
F:200,201/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 91.7%; Score 812.5; DB 2; Length 201;  
Best Local Similarity 81.2%; Pred. No. 1.5e-58;  
Matches 164; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

OY 4 MNPEYDYLFKLLIGDSGVKSCILLRFADDTYESYSTIGVDFKRTIELDGKTI 63  
DB 1 MNPEYDYLFKLLIGDSGVKSCILLRFADDTYESYSTIGVDFKRTIELDGKTI 60  
OY 64 I-----ESFNNVKOMLOEIDRYASENVNKL 91  
DB 61 IWDTAGQERFRRTTSYRGAGIIIVYDVDOESFNNVKOMLOEIDRYASENVNKL 120  
OY 92 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRKGPGA 151  
DB 121 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRKGPGA 180

DB 121 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRKGPGA 179  
OY 152 GAESNVKIOSTPVKSGGCGC 173  
DB 180 GSEKSNVNIQSTPVKSGGCGC 201

RESULT 6  
S06147  
GTP-binding protein rab1b - rat  
M:Alternate names: ras-related protein rab1b  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 02-Feb-2001  
C:Accession: S06147; S03189  
R:Touchot, N.; Zahraoui, A.; Vialh, E.; Tavilian, A.  
FEBS Lett. 236, 79-84, 1989  
A:Title: Biochemical properties of the YPT-related rab1b protein. Comparison with rab  
A:Reference number: S06147; MUID:90033316; PMID:2509243  
A:Accession: S06147  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-201 <TUD>  
R:Vielh, E.; Touchot, N.; Zahraoui, A.; Tavilian, A.  
Nucleic Acids Res. 17, 1770, 1989  
A:Title: Nucleotide sequence of a rat cDNA: RAB1B, encoding a RAB1-YPT related protel  
A:Reference number: S03189; MUID:89160341; PMID:2493636  
A:Accession: S03189  
A:Molecule type: mRNA  
A:Residues: 1-95, 'YA', '98-201 <VIE>  
A:Cross-references: EMBL:X13905; MUID:957005; PIDN:CAA32105.1; PID:957006  
C:Genetics:  
A:Gene: rab1B  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop  
F:9-124/Domain: translation elongation factor Tu homology <ETU>  
F:15-22/Region: nucleotide-binding motif A (P-loop)  
F:121-124/Region: GTP-binding NKXD motif  
F:151-153/Region: GTP-binding SAK/L motif  
F:21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asp, Lys, Asp, Ser) #  
F:200,201/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 87.1%; Score 771.5; DB 2; Length 201;  
Best Local Similarity 77.2%; Pred. No. 3e-55;  
Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2;

OY 4 MNPEYDYLFKLLIGDSGVKSCILLRFADDTYESYSTIGVDFKRTIELDGKTI 63  
DB 1 MNPEYDYLFKLLIGDSGVKSCILLRFADDTYESYSTIGVDFKRTIELDGKTI 60  
OY 64 I-----ESFNNVKOMLOEIDRYASENVNKL 91  
DB 61 IWDTAGQERFRRTTSYRGAGIIIVYDVDOESFNNVKOMLOEIDRYASENVNKL 120  
OY 92 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRKGPGA 151  
DB 121 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRKGPGA 180  
OY 152 GAESNVKIOSTPVKSGGCGC 173  
DB 181 G-ERPMLKIDSTPVKSGGCGC 201

RESULT 7  
S38339  
GTP-binding protein rab1 - great pond snail  
C:Species: Lymnaea stagnalis (great pond snail)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 02-Feb-2001  
C:Accession: S38339; S32206  
R:Atterberg, M.; van Die, I.; Yang, H.; Andriessen, J.A.; van Tetering, A.; van den E  
Eur. J. Biochem. 217, 241-246, 1993  
A:Title: Isolation and characterization of three cDNAs coding for Rab proteins from t  
A:Reference number: S38339; MUID:94039042; PMID:8223561  
A:Accession: S38339

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-205 <AGT>  
A:Cross-references: EMBL:X72688; NID:g288933; PIDN:CAA51233.1; PID:g288934  
C:Genetics:  
A:Gene: rab1  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine  
F:12-127/Domain: translation elongation factor Tu homology <ETD>  
F:18-25/Region: nucleotide-binding motif A (P-loop)  
F:124-127/Region: GTP-binding NKXD motif  
F:154-156/Region: GTP-binding SAR/L motif  
F:24,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta  
F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 82.7%; Score 733; DB 2; Length 205;  
Best Local Similarity 73.8%; Pred. No. 4e-52;  
Matches 152; Conservative 7; Mismatches 13; Indels 34; Gaps 3;

QY 1 MSANPEYDYLKLLIGSGVKSCLLRPADDTYESYSTIGVDFKIRITELDGKTI 60  
DB 1 MSTMPDYDLKLLIGSGVKSCLLRPADDTYESYSTIGVDFKIRITELDGKTI 60  
QY 61 KLOI-----ESFNWKQWLOEIDRYASENVNL 88  
DB 61 KLOIMDTAGQERFRITTSYRGAGIIVYDVTDOESFNWKQWLOEIDRYASENVNL 120  
QY 89 LVGNKCDLTKKRVYDTTAKERADSLGIPLETSAKNAATNVQSFMTAAEIKRRGPGA 148  
DB 121 LVGNKSDLTTRKRVYDPTTAKERADSLGIPLETSAKNAATNVQSFMTAAEIKRRGPGA-1 179  
QY 149 TAGGAESNVKTI-QSTPVKQSGGCC 173  
DB 180 TAASDSKPSVKINSSTPVSAKNGGCC 205

RESULT 8  
JEO318  
GTP-binding protein rabB - silkworm  
C:Species: Bombyx mori (silkworm)  
C>Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 02-Feb-2001  
C:Accession: JEO318  
R:Uno, T.; Ueno, M.; Nakajima, A.; Shirai, Y.; Alzono, Y.  
Biosci. Biotechnol. Biochem. 62, 1885-1891, 1998  
A:Title: Molecular cloning of cDNA for Rab from the brain of Bombyx mori and biochemica  
A:Reference number: JEO318; M0ID:99053143; PMID:9836423  
A:Accession: JEO318  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-202 <UNO>  
A:Cross-references: GB:AF013572  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:9-124/Domain: translation elongation factor Tu homology <ETD>  
F:15-22/Region: nucleotide-binding motif A (P-loop)  
F:121-124/Region: GTP-binding NKXD motif  
F:151-153/Region: GTP-binding SAR/L motif  
F:21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 75.7%; Score 671; DB 2; Length 202;  
Best Local Similarity 66.5%; Pred. No. 4e-47;  
Matches 135; Conservative 18; Mismatches 16; Indels 34; Gaps 3;

QY 4 MNPEDYLFKLLIGSGVKSCLLRPADDTYESYSTIGVDFKIRITELDGKTIKIQ 63  
DB 1 MNPEDYLFKLLIGSGVKSCLLRPADDTYESYSTIGVDFKIRITELDGKTIKIQ 60  
QY 64 I-----ESFNWKQWLOEIDRYASENVNLKLVG 91  
DB 61 IMDTAGQERFRITTSYRGAGIIVYDVTDOESFNWKQWLOEIDRYASENVNLKLVG 120  
QY 92 NKCDLTKKRVYDTTAKERADSLGIPLETSAKNAATNVQSFMTAAEIKRRGPGA 151  
DB 92 NKCDLTKKRVYDTTAKERADSLGIPLETSAKNAATNVQSFMTAAEIKRRGPGA 151

DB 121 NKCDLTKKRVYDSTAKOYAEOLGIPLETSAKNSTNVQAFMTAAEIKRARGPST-G 179  
QY 152 GAESNVKTI-QSTPVKQSGGCC 173  
DB 180 MAPAGHKIDGQGPIDTGKSSCC 202

RESULT 9  
T33781  
hypothetical protein C39F7.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 11-Jan-2000  
C:Accession: T33781  
R:Magill, L.; Scheet, P.; Dubbelde, C.  
submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of C. elegans cosmid C39F7.  
A:Reference number: Z21407  
A:Accession: T33781  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-205 <MAG>  
A:Cross-references: EMBL:AF101310; PIDN:AC69218.1; GSPDB:GN00023; CESP:C39F7.4  
A:Experimental source: strain Bristol N2; clone C39F7  
C:Genetics:  
A:Gene: CESP:C39F7.4  
A:Map position: 5  
A:Insertions: 8/2; 48/3; 64/3; 96/3; 163/3  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 74.5%; Score 660; DB 2; Length 205;  
Best Local Similarity 66.2%; Pred. No. 3.2e-46;  
Matches 137; Conservative 16; Mismatches 18; Indels 36; Gaps 4;

QY 1 MSANPEYDYLKLLIGSGVKSCLLRPADDTYESYSTIGVDFKIRITELDGKTI 60  
DB 1 MAANPEYDYLKLLIGSGVKSCLLRPADDTYESYSTIGVDFKIRITELDGKTI 60  
QY 61 KLOI-----ESFNWKQWLOEIDRYASENVNL 88  
DB 61 KLOIMDTAGQERFRITTSYRGAGIIVYDVTDOESFNWKQWLOEIDRYASENVNL 120  
QY 89 LVGNKCDLTKKRVYDTTAKERADSLGIPLETSAKNAATNVQSFMTAAEIKRRGPGA 148  
DB 121 LVGNKCDLTKKRVYDTTAKERADSLGIPLETSAKNAATNVQAFMTAAEIKRRGPGA 180  
QY 149 TAGGAESNVKTI-QSTPVKQSGGCC 173  
DB 181 GAGGA--PEVRLTGSGPVDKRGSGCC 205

RESULT 10  
S30096  
GTP-binding protein ypt1 (similarity) - Neurospora crassa  
N:Alternate names: protein B9J10.240; ras-related protein ypt1  
C:Species: Neurospora crassa  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 19-Jan-2001  
C:Accession: S30096; T49787  
R:Beilutz, K.; Palme, K.; Dieffenhal, T.; Russo, V.E.A.  
Mol. Gen. Genet. 235, 413-421, 1992  
A:Title: The Ncyp1 gene from Neurospora crassa is located on chromosome 2: molecular  
A:Reference number: S30096; M0ID:93101148; PMID:1361212  
A:Accession: S30096  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-203 <HEI>  
A:Cross-references: GB:S51252; NID:g262045; PIDN:AA824564.1; PID:g262046  
R:Schulte, U.; Align, V.; Honeisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatu  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022  
A:Accession: T49787  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-203 <SCH>

A:Cross-references: EMBL:AL356324; GSPDB:GN00116; NCSP:B9J10.240  
 A:Experimental source: BAC clone B9J10; strain OR74A  
 C:Genetics:  
 A:Gene: ypt1; NCSP:B9J10.240  
 A:Map position: 6  
 A:Introns: 4/2; 45/3; 61/3; 137/3  
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
 C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine  
 F:9-124/Domain: translation elongation factor Tu homology <ETU>  
 F:15-22/Region: nucleotide-binding motif A (P-loop)  
 F:121-124/Region: GTP-binding SAK/L motif  
 F:151-153/Region: GTP-binding SAK/L motif  
 F:21-22/Region: GTP-binding SAK/L motif  
 F:21-22/Region: GTP-binding SAK/L motif  
 F:202/203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 74.3%; Score 658.5; DB 2; Length 203;  
 Best Local Similarity 66.5%; Pred. No. 4,2e-46;  
 Matches 135; Conservative 11; Mismatches 24; Indels 33; Gaps 2;

OY 4 MNPEYDLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDKRTIKLQ 63  
 |||||  
 DB 1 MNPEYDLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDKRTIKLQ 60  
 |||||

OY 64 I-----ESFNNVKOMLQEDIRYASENVNKLIVG 91  
 |||||

DB 61 IMPTAGQERFRTITSSYRGANGIIIVDVTDDSFNNVKOMLQEDIRYASENVNKLIVG 120  
 |||||

OY 92 NKCDLTKKVVDTYAKFEADSLGIPLETSAKNATVQSPMTAAETKRKMGPGATAG 151  
 |||||

DB 121 NKSDMDKRVYEVYAKFEADSLGIPLETSAKNATVQSPMTAAETKRKMGPGATAG 180  
 |||||

OY 152 GAERKSNVKIQTSPYKOSGGGCC 173  
 |||||

DB 181 NTKASVNVSRGCHVSNSSGCC 203  
 |||||

RESULT 11  
 JC4105  
 GTP-binding protein yptc1 - Chlamydomonas reinhardtii  
 A:Alternate names: membrane vesicle transport protein yptc1; ras-like yptc1 protein; smc  
 C:Species: Chlamydomonas reinhardtii  
 C:Date: 21-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 02-Feb-2001  
 C:Accession: J04105  
 R:Dieterle, W.; Fabry, S.; Huber, H.; Schmitt, R.  
 Gene 158, 41-50, 1995  
 A:Title: Analysis of a family of ypt genes and their products from Chlamydomonas reinhardtii  
 A:Reference number: J04105; MUID:95309723; PMID:7789809  
 A:Accession: J04105  
 A:Molecule type: mRNA  
 A:Residues: 1-203 <ID>  
 A:Cross-references: GB:U13168; NID:9806723; PID:AAA82727.1; PID:9806724  
 C:Comment: This protein plays an essential role in the regulation of intracellular membrane traffic  
 C:Genetics:  
 A:Gene: yptc1  
 A:Introns: 5/1; 10/2; 29/2; 45/2; 61/2; 85/2; 137/2; 172/1  
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
 C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
 F:9-124/Domain: translation elongation factor Tu homology <ETU>  
 F:15-22/Region: nucleotide-binding motif A (P-loop)  
 F:121-124/Region: GTP-binding SAK/L motif  
 F:151-153/Region: GTP-binding SAK/L motif  
 F:21-22/Region: GTP-binding SAK/L motif  
 F:21-22/Region: GTP-binding SAK/L motif  
 F:202/203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 73.0%; Score 646.5; DB 2; Length 203;  
 Best Local Similarity 65.7%; Pred. No. 3,9e-45;  
 Matches 134; Conservative 9; Mismatches 26; Indels 35; Gaps 3;

OY 4 MNPEYDLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDKRTIKLQ 63  
 |||||

DB 1 MNPEYDLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDKRTIKLQ 60  
 |||||

OY 64 I-----ESFNNVKOMLQEDIRYASENVNKLIVG 91  
 |||||

DB 61 IMPTAGQERFRTITSSYRGANGIIIVDVTDDSFNNVKOMLQEDIRYASENVNKLIVG 120  
 |||||

OY 92 NKCDLTKKVVDTYAKFEADSLGIPLETSAKNATVQSPMTAAETKRKMGPGATAG 151  
 |||||

DB 121 NKSDMDKRVYEVYAKFEADSLGIPLETSAKNATVQSPMTAAETKRKMGPGATAG 180  
 |||||

OY 150 AGAERKSNVKIQTSPYKOSGGGCC 173  
 |||||

DB 181 KAGGPVVRQ-EGKPIKNSSSCC 203  
 |||||

RESULT 12  
 T50323  
 ypt1-related protein 1 [imported] - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 02-Sep-2000  
 C:Accession: T50323  
 R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Cadieu, E.; Lelaure, V.; Gilbert  
 submitted to the EMBL Data Library, January 2000  
 A:Reference number: 225061  
 A:Accession: T50323  
 A:Status: preliminary; translated from GB/EMBL/DDBB  
 A:Molecule type: DNA  
 A:Residues: 1-203 <MCD>  
 A:Cross-references: EMBL:AL356324; GSPDB:GN00067; SPDB:SPBC1703.10  
 C:Genetics:  
 A:Gene: SPDB:SPBC1703.10  
 A:Map position: 2  
 A:Introns: 4/2; 45/3; 61/3; 94/1  
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
 F:9-124/Domain: translation elongation factor Tu homology <ETU>

Query Match 73.0%; Score 646.5; DB 2; Length 203;  
 Best Local Similarity 64.5%; Pred. No. 3,9e-45;  
 Matches 131; Conservative 16; Mismatches 23; Indels 33; Gaps 2;

OY 4 MNPEYDLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDKRTIKLQ 63  
 |||||

DB 1 MNPEYDLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDKRTIKLQ 60  
 |||||

OY 64 I-----ESFNNVKOMLQEDIRYASENVNKLIVG 91  
 |||||

DB 61 IMPTAGQERFRTITSSYRGANGIIIVDVTDDSFNNVKOMLQEDIRYASENVNKLIVG 120  
 |||||

OY 92 NKCDLTKKVVDTYAKFEADSLGIPLETSAKNATVQSPMTAAETKRKMGPGATAG 151  
 |||||

DB 121 NKSDMDKRVYEVYAKFEADSLGIPLETSAKNATVQSPMTAAETKRKMGPGATAG 180  
 |||||

OY 152 GAERKSNVKIQTSPYKOSGGGCC 173  
 |||||

DB 181 SNAKSVNRVGGGTNVSQSSGCC 203  
 |||||

RESULT 13  
 JC1247  
 GTP-binding protein yptv1 - Volvox carterii  
 C:Species: Volvox carterii  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 02-Feb-2001  
 C:Accession: J01247  
 R:Fabry, S.; Nat, N.; Huber, H.; Palme, K.; Jaenicke, L.; Schmitt, R.  
 Gene 116, 153-162, 1992  
 A:Title: The yptv1 gene encodes a small G-protein in the green alga Volvox carterii: G  
 A:Reference number: J01247; MUID:92280499; PMID:1511889  
 A:Accession: J01247  
 A:Molecule type: DNA  
 A:Residues: 1-203 <FAB>  
 A:Cross-references: GB:M3438; NID:9170660; PID:AAA4255.1; PID:9170661  
 C:Genetics:  
 A:Gene: yptv1  
 A:Introns: 5/2; 10/3; 29/3; 45/3; 61/3; 85/3; 137/3; 172/2



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:15:45 ; Search time 22 Seconds  
(without alignments)  
326.155 Million cell updates/sec

Title: US-09-820-003b-2

Perfect score: 886  
Sequence: 1 MSSWNPEDYLFKLLIGDS.....EKSNVKIQSTPVMQSGGCGC 173

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	860	97.1	205	1	RB1A_HUMAN
2	849	95.8	205	1	RB1A_RAT
3	828	93.5	202	1	RB1A_DISOM
4	760.5	85.8	201	1	RB1B_RAT
5	733	82.7	205	1	RB1A_LYMST
6	658.5	74.3	203	1	YPT1_NEUCR
7	646.5	73.0	203	1	YPT1_CHURE
8	646.5	73.0	203	1	YPT1_SCHPO
9	642.5	72.5	203	1	YPT1_VOICA
10	631.5	71.3	201	1	YPT1_PHTIN
11	619.5	69.9	208	1	ARAB_AKATH
12	618.5	69.8	203	1	YPT2_MAIZE
13	613.5	69.2	202	1	RIC1_ORISA
14	571	64.4	206	1	YPT1_YEAST
15	571	64.4	206	1	YPT1_YEAST
16	538	61.2	199	1	RB1A_DICDI
17	523	59.0	208	1	YPT1_DICDI
18	468.5	52.9	208	1	SAS1_DICDI
19	461	52.0	203	1	SAS2_DICDI
20	442	49.9	215	1	RAV1_BERVU
21	438.5	49.5	216	1	ARAB_AKATH
22	435.5	49.2	210	1	YPT2_SCHPO
23	424	47.9	210	1	RBAB_DISOM
24	418	47.2	207	1	RBAB_HUMAN
25	417	47.1	207	1	RBAB_RAT
26	414.5	46.8	199	1	RBAB_DICDI
27	410.5	46.7	207	1	RBAB_HUMAN
28	405.5	45.8	206	1	YPT2_VOICA
29	405.5	45.8	200	1	RB10_HUMAN
30	402.5	45.4	200	1	RB10_CANEA
31	398	44.9	200	1	RAO1_DISOM
32	398	44.9	200	1	RAO1_DISOM
33	398	44.9	203	1	RYL1_YARLI

34	384.5	43.4	203	1	RB13_HUMAN
35	374.5	42.3	200	1	RB10_RAT
36	373.5	42.2	215	1	SECA_YEAST
37	360.5	40.7	201	1	RB35_HUMAN
38	360	40.6	210	1	SECA_CANAL
39	352	39.7	209	1	RB2A_MAIZE
40	349.5	39.4	210	1	RB2B_MAIZE
41	344.5	38.9	213	1	YPT4_CHURE
42	342	38.6	219	1	RB3A_BOVIN
43	339	38.3	213	1	YPT4_VOICA
44	336	37.9	220	1	RB3A_DROME
45	336	37.9	220	1	RB3A_HUMAN

## ALIGNMENTS

RESULT 1	ID	RB1A_HUMAN	STANDARD:	PRT:	205 AA.
AC	P1476:				
DT	01-OCT-1989	(Rel. 12, Created)			
DT	01-OCT-1989	(Rel. 12, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Ras-related protein Rab-1A (YPT1-related protein).				
GN	RB1 OR RB1A.				
OS	Homo sapiens (Human), and				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID:9606, 10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=Human;				
RA	MEDLINE=89308668; PubMed=2501306;				
RA	Zahraoui A., Touchot N., Chardin P., Tavilian A.;				
RT	"The human Rab genes encode a family of GTP-binding proteins related				
RT	to yeast YPT1 and SEC4 products involved in secretion.";				
RL	J. Biol. Chem. 264:12394-12401(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=Human; TISSUE=Placenta;				
RA	Strausberg R.;				
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=Mouse;				
RA	MEDLINE=8816649; PubMed=3127202;				
RA	Haudruck H., Disela C., Wagner P., Gallwitz D.;				
RT	"The ras-related ypt protein is an ubiquitous eukaryotic protein:				
RT	isolation and sequence analysis of mouse cDNA clones highly				
RT	homologous to the yeast YPT1 gene.";				
RL	EMBO J. 6:4049-4053(1987).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=Mouse;				
RA	MEDLINE=89386011; PubMed=2506528;				
RA	Wichmann H., Disela C., Haudruck H., Gallwitz D.;				
RT	"Nucleotide sequence of the mouse ypt1 gene encoding a ras-related				
RT	GTP-binding protein.";				
RT	Nucleic Acids Res. 17:6737-6738(1989).				
RL	[5]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=Mouse;				
RA	Wu G., Dorn G.W. II;				
RT	"Mouse Rab1a, member of Ras oncogene family mRNA.";				
RT	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	PHOSPHORYLATION BY CDC2.				
RA	MEDLINE=9121885; PubMed=1902553;				
RA	Bally E., McGeefey M., Touchot N., Zahraoui A., Goud B., Bornes M.;				
RT	"Phosphorylation of two small GTP-binding proteins of the Rab family				
RT	by p34cdc2.";				

RL Nature 350:715-718(1991).  
 RN [7]  
 RP ISOPRENOL.  
 RX MEDLINE-91296801; PubMed-1648736;  
 RA Khosravi-Far R., Lutz R.J., Cox A.D., Conroy L., Bourne J.R.,  
 RT Sinensky M., Balch W.E., Buss J.E., Der C.J.,  
 "Isoprenoid modification of rab proteins terminating in CC or CXC motifs."  
 RT Proc. Natl. Acad. Sci. U.S.A. 88:6264-6268(1991).  
 CC -1- FUNCTION: PROBABLY REQUIRED FOR TRANSIT OF PROTEIN FROM THE ER  
 CC THROUGH GOLGI COMPARTMENT.  
 CC -1- SUBCELLULAR LOCATION: GOLGI COMPARTMENT.  
 CC -1- PTM: PHOSPHORYLATED BY CDC2 KINASE DURING MITOSIS.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M28209; AAA60240.1; -  
 DR EMBL: BC000905; AAH00905.1; -  
 DR EMBL: Y00094; CAA68284.1; -  
 DR EMBL: X15744; CAA33760.1; -  
 DR EMBL: X15745; CAA33760.1; JOINED.  
 DR EMBL: X15746; CAA33760.1; JOINED.  
 DR EMBL: X15747; CAA33760.1; JOINED.  
 DR EMBL: AF226873; AAF33844.1; -  
 DR PIR: S05551; TWSTP.  
 DR PIR: A34323; A34323.  
 DR PIR: S19104; S19104.  
 DR HSSP: P05713; 3RAB.  
 DR PMMA-2DPAGE: P11476; -  
 DR GeneW: HGNC:9758; RAB1A.  
 DR MIM: 179508; -  
 DR MGI: 97842; Rab1.  
 DR InterPro: IPR003579; GTPase\_Rab.  
 DR InterPro: IPR001806; Ras\_trnsfmg.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF00071; ras; 1.  
 DR PRINTS: PR00449; RASTRNSFRMG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFAMs: TIGR00231; small\_GTP; 1.  
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack;  
 KW Endoplasmic reticulum; Phosphorylation.  
 FT NP\_BIND 18 25 GTP (BY SIMILARITY).  
 FT NP\_BIND 66 70 GTP (BY SIMILARITY).  
 FT NP\_BIND 124 127 GTP (BY SIMILARITY).  
 FT DOMAIN 40 48 EFFECTOR REGION (BY SIMILARITY).  
 FT MOD\_RES 194 194 PHOSPHORYLATION (BY CDC2) (PROBABLE).  
 FT LIPID 204 204 GERANYL-GERANYL.  
 FT LIPID 205 205 GERANYL-GERANYL.  
 SQ SEQUENCE 205 AA: 22678 MW: 22678 MW: B2A8F43B0FB17D6 CRC64:  
 Query Match 97.1%; Score 860; DB 1; Length 205;  
 Best Local Similarity 84.4%; Pred. No. 9, 3e-65;  
 Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;  
 QY 1 MSSANPEHYLRLKLLIGSGYKSCLLLRFPDDITYESTISTIGVDFKIRTELDGKTI 60  
 DB 1 MSSANPEHYLRLKLLIGSGYKSCLLLRFPDDITYESTISTIGVDFKIRTELDGKTI 60  
 QY 61 KLOI-----ESFNNTKOWLOEIDRASEVNTL 88  
 DB 61 KLOIINDTAGOERFRITTSYNGAGCIIVYVDYDSEFNNTKOWLOEIDRASEVNTL 120  
 QY 89 LVGNACDLTKKRVYDTAKERFADSLGIPFLETSANATNVDSFMTAAETIKRKGPA 148  
 DB 121 LVGNACDLTKKRVYDTAKERFADSLGIPFLETSANATNVDSFMTAAETIKRKGPA 180

QY 149 TAGAERKSNVKTIOSTFVKSGGCC 173  
 DB 181 TAGAERKSNVKTIOSTFVKSGGCC 205  
 RESULT 2  
 ID RBLA\_RAT STANDARD: PRT; 205 AA.  
 AC P05711;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ras-related protein Rab-1A.  
 GN RAB1 OR RAB1A.  
 OS Rattus norvegicus (Rat), and  
 OC Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116, 9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Rat;  
 RX MEDLINE-88068563; PubMed-3317403;  
 RA Touchot N., Chardin P., Tavilian A.;  
 RT Four additional members of the ras gene superfamily isolated by an  
 RT oligonucleotide strategy: molecular cloning of YPT-related cDNAs from  
 RT a rat brain library."  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8210-8214(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-C.familiaris; STRAIN-Cocker spaniel;  
 RX MEDLINE-91061765; PubMed-2123294;  
 RA Chavrier P., Vangron M., Sander C., Simons K., Zerial M.;  
 RT "Molecular cloning of YP11/SEC4-related cDNAs from an epithelial cell  
 RT line."  
 RL Mol. Cell. Biol. 10:6578-6585(1990).  
 RN [3]  
 RP ISOPRENOL.  
 RX MEDLINE-91296801; PubMed-1648736;  
 RA Khosravi-Far R., Lutz R.J., Cox A.D., Conroy L., Bourne J.R.,  
 RT Sinensky M., Balch W.E., Buss J.E., Der C.J.;  
 RT "Isoprenoid modification of rab proteins terminating in CC or CXC motifs."  
 RT Proc. Natl. Acad. Sci. U.S.A. 88:6264-6268(1991).  
 CC -1- FUNCTION: PROBABLY REQUIRED FOR TRANSIT OF PROTEIN FROM THE ER  
 CC THROUGH GOLGI COMPARTMENT.  
 CC -1- MISCELLANEOUS: RAB-1A BINDS GTP AND GDP AND POSSESS INTRINSIC  
 CC GTPASE ACTIVITY.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: J02998; AAA42006.1; -  
 DR EMBL: X56384; CAA56775.1; ALT\_SEQ.  
 DR PIR: A36364; A36364.  
 DR PIR: A39963; A39963.  
 DR HSSP: P05713; 3RAB.  
 DR InterPro: IPR003579; GTPase\_Rab.  
 DR InterPro: IPR001806; Ras\_trnsfmg.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF00071; ras; 1.  
 DR PRINTS: PR00449; RASTRNSFRMG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFAMs: TIGR00231; small\_GTP; 1.  
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack;  
 KW Endoplasmic reticulum.  
 FT NP\_BIND 18 25 GTP (BY SIMILARITY).

```

FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 124 127 GTP (BY SIMILARITY).
FT DOMAIN 40 48 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL.
FT LIPID 205 205 GERANYL-GERANYL.
SQ SEQUENCE 205 AA; 22763 MW; 34649523B0FB0E6FC CRC64;

Query Match 95.8%; Score 849; DB 1; Length 205;
Best Local Similarity 83.4%; Pred. No. 7.7e-64;
Matches 171; Conservative 0; Mismatches 2; Indels 32; Gaps 1;

QY 1 MSNAPDYELFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDFKRIEIDGKTI 60
DB 1 MSNAPDYELFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDFKRIEIDGKTI 60
QY 61 KIQI-----ESFNVKOMLOEIDRYASENVNKL 88
DB 61 KIQIMDTAGGERFRTTSSYRGAGIIVYDVTDQESFNVKOMLOEIDRYASENVNKL 120
QY 89 LVGKCDLTKKVVDTTAKFADSLGIPLETSAKNATVROSPMTMAEIKKMGPGA 148
DB 121 LVGKCDLTKKVVDTTAKFADSLGIPLETSAKNATVROSPMTMAEIKKMGPGA 180
QY 149 TAGAEKSNVKIOSPPVKGSGGCC 173
DB 181 TAGAEKSNVKIOSPPVKGSGGCC 205

RESULT 3
RAB1_DISOM
ID RAB1_DISOM STANDARD: PRT: 202 AA.
AC P2125;
DR 01-AUG-1991 (Rel. 19, Created)
DR 01-AUG-1991 (Rel. 19, Last sequence update)
DR 01-OCT-1994 (Rel. 30, Last annotation update)
DE Ras-related protein ORAB-1.
OS Discospora ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hypnosqualea; Pristigaster; Batoidae;
OC Torpediniformes; Narcinidae; Narcinidae; Discopidae.
OX NCBI_Taxid=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Electric lobe;
RA MEDLINE-91115900; PubMed-1899244;
RA Ngeese J.K., Elferink L.A., Scheller R.H.;
RT "A family of ras-like GTP-binding proteins expressed in electrosensor
neurons."
RL J. Biol. Chem. 266:2675-2680(1991).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
TRAFFIC (BY SIMILARITY).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch)
CC EMBL: M38393; AAA49234.1;
DR PIR: D38625; D38625.
DR HSP: P05713; 3A88.
DR InterPro: IPR003578; GTPase_Rab.
DR InterPro: IPR001806; Ras_transfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras_1.
DR PRINTS: PR00449; RASTNSPRMG.
DR SMART: SM00175; Ras_1.
DR TIGRfams: TIGR00231; small_GTP_1.
KW GTP binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).

```

```

FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 202 202 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 202 AA; 22333 MW; 31078502BCDD1B9 CRC64;

Query Match 93.5%; Score 828; DB 1; Length 202;
Best Local Similarity 81.7%; Pred. No. 4.2e-62;
Matches 165; Conservative 3; Mismatches 2; Indels 32; Gaps 1;

QY 4 MNPEVYELFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDFKRIEIDGKTI 63
DB 1 MNPEVYELFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDFKRIEIDGKTI 60
QY 64 I-----ESFNVKOMLOEIDRYASENVNKL 91
DB 61 IMPTAGGERFRTTSSYRGAGIIVYDVTDQESFNVKOMLOEIDRYASENVNKL 120
QY 92 NKCDLTKKVVDTTAKFADSLGIPLETSAKNATVROSPMTMAEIKKMGPGA 151
DB 121 NKCDLTKKVVDTTAKFADSLGIPLETSAKNATVROSPMTMAEIKKMGPGA 180
QY 152 GAEKSNVKIOSPPVKGSGGCC 173
DB 181 GAEKSNVKIOSPPVKGSGGCC 202

RESULT 4
RAB1_RAT
ID RAB1_RAT STANDARD: PRT: 201 AA.
AC P10536;
DR 01-JUL-1989 (Rel. 11, Created)
DR 01-JUL-1989 (Rel. 11, Last sequence update)
DR 01-OCT-1994 (Rel. 30, Last annotation update)
DE Ras-related protein Rab-1b.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-89160341; PubMed-2493636;
RA Zahraoui A., Touchot N., Chardin P., Tavilian A.;
RT "Nucleotide sequence of a rat cDNA: rab1b, encoding a rab1-YPT
related protein."
RL Nucleic Acids Res. 17:1770-1770(1989).
RN [2]
RP CHARACTERIZATION.
RC MEDLINE-90033316; PubMed-2509243;
RA Touchot N., Zahraoui A., Vielh E.,
RT "Biochemical properties of the YPT-related rab1b protein. Comparison
with rab1a."
RL FEBS Lett. 256:79-84(1989).
RN [3]
RP ISOPRENOID.
RC MEDLINE-91296801; PubMed-1648736;
RA Khosravi-Far R., Lutz R.J., Cox A.D., Conroy L., Bourne J.R.,
RA Stenseth M., Balch W.E., Buss J.E., Der C.J.;
RT "Isoprenoid modification of rab proteins terminating in CC or CXC
motifs."
RL Proc. Natl. Acad. Sci. U.S.A. 88:6264-6268(1991).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
TRAFFIC (BY SIMILARITY).
CC -1- MISCELLANEOUS: RAB-1B BINDS GTP AND GDP AND POSSESSES INTRINSIC
GTPASE ACTIVITY.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

```



modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X13905; CAA32105.1; -

DR PIR: S03189; S03189.

DR PIR: S06147; S06147.

DR HSSP: P05713; 3RAB.

DR InterPro: IPR003579; GTPase\_Rab.

DR InterPro: IPR001806; Ras\_transfmg.

DR InterPro: IPR005225; Small\_GTP.

DR Pfam: PF00071; ras; 1.

DR PRINTS: PRO0449; RASTRNSFRNG.

DR SMART: SM00175; RAB; 1.

DR TIGRfams: TIGR00231; small\_gtp; 1.

KW GTP-binding; Lipoprotein; Prenylation; Protein transport.

FT NP\_BIND 15 22 GTP (BY SIMILARITY).

FT NP\_BIND 63 67 GTP (BY SIMILARITY).

FT NP\_BIND 121 124 GTP (BY SIMILARITY).

FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).

FT LIPID 200 200 GERANYL-GERANYL.

FT LIPID 201 201 GERANYL-GERANYL.

FT MUTAGEN 21 21 K->M: ABOLISHES GTP-BINDING.

SO SEQUENCE 201 AA; 22163 MW; 8D3EECC2AEFA42FE CRC64;

Query Match 85.8%; Score 760.5; DB 1; Length 201;

Best Local Similarity 76.2%; Pred. No. 1.7e-56;

Matches 154; Conservative 7; Mismatches 8; Indels 33; Gaps 2;

OY 4 MNPEVDYLFKLLIGDSGVGSKCLLRADDTYESTYSTIGVDKRTIELDKGTI 63

DB 1 MNPEVDYLFKLLIGDSGVGSKCLLRADDTYESTYSTIGVDKRTIELDKGTI 60

OY 64 I-----ESFNVKQWLQEIIDRYASENVKLLVG 91

DB 61 IWDVAGREFRFTVSSYRGANGIIIVYDVTDQESFNVKQWLQEIIDRYASENVKLLVG 120

OY 92 MKCOLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGANG 151

DB 121 MKSDLTTRKVVNDYTTAKFADSLGIPLETSAKNATNVEOAFMTAAEIKRMGPGANG 180

OY 152 GAESKNVRIKOSTPYKSGGGCC 173

DB 181 GERPNLKLIDSTPYKSGGGCC 201

RESULT 5

RAB1\_LYMS

ID RAB1\_LYMS STANDARD; PRT; 205 AA.

AC 005974;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Ras-related protein Rab-1A.

GN RAB1A.

OS Lymnaea stagnalis (Great pond snail).

CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;

CC Lymnaeidae; Lymnaea.

CC NCBI\_TaxID=6523;

OX NCBI\_TaxID=6523;

RN [1]

RP MEDLINE FROM N.A.

RX MEDLINE-94039042; PubMed-8223561;

RA Agterberg M., van Die I., Yang H., Andriessen J.A.,

RA van Tetering A., van den Eljnden D.H., Ploegh H.L.;

RT "Isolation and characterization of three cDNAs coding for Rab

RT proteins from the albumen gland of the mollusc Lymnaea stagnalis.";

RL Eur. J. Biochem. 217:241-246(1993).

CC -1- FUNCTION: PROBABLY REQUIRED FOR TRANSIT OF PROTEIN FROM THE ER

CC THROUGH GOLGI COMPARTMENT (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: GOLGI COMPARTMENT (BY SIMILARITY).

CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X72688; CAA51233.1; -

DR PIR: S32206; S32206.

DR PIR: S38339; S38339.

DR HSSP: P05713; 3RAB.

DR InterPro: IPR003579; GTPase\_Rab.

DR InterPro: IPR001806; Ras\_transfmg.

DR InterPro: IPR005225; Small\_GTP.

DR Pfam: PF00071; ras; 1.

DR PRINTS: PRO0449; RASTRNSFRNG.

DR SMART: SM00175; RAB; 1.

DR TIGRfams: TIGR00231; small\_gtp; 1.

KW GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack;

FT NP\_BIND 18 25 GTP (BY SIMILARITY).

FT NP\_BIND 66 70 GTP (BY SIMILARITY).

FT NP\_BIND 124 127 GTP (BY SIMILARITY).

FT DOMAIN 40 48 EFFECTOR REGION (BY SIMILARITY).

FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).

FT LIPID 205 205 GERANYL-GERANYL (BY SIMILARITY).

SO SEQUENCE 205 AA; 22760 MW; D5E0BDCB84A26FE CRC64;

Query Match 82.7%; Score 733; DB 1; Length 205;

Best Local Similarity 73.8%; Pred. No. 3.4e-54;

Matches 152; Conservative 7; Mismatches 13; Indels 34; Gaps 3;

OY 1 MSSNPEVDYLFKLLIGDSGVGSKCLLRADDTYESTYSTIGVDKRTIELDKGTI 60

DB 1 MSTNPEVDYLFKLLIGDSGVGSKCLLRADDTYESTYSTIGVDKRTIELDKGTI 60

OY 61 KIQI-----ESFNVKQWLQEIIDRYASENVKLL 88

DB 61 KIQIWDVAGREFRFTVSSYRGANGIIIVYDVTDQESFNVKQWLQEIIDRYASENVKLL 120

OY 89 LVGNKCDLTTRKVVNDYTTAKFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGANG 148

DB 121 LVGNKSDLTTRKVVNDYTTAKFADSLGIPLETSAKNATNVEOAFMTAAEIKRMGPGANG 179

OY 149 TAGGAESKNVRIKOSTPYKSGGGCC 173

DB 180 TAASDKPSVKINSTPYKSGGGCC 205

RESULT 6

YPT1\_NEUCR

ID YPT1\_NEUCR STANDARD; PRT; 203 AA.

AC P33723;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE GTP-binding protein ypt1.

GN YPT-1 OR B9J10.240.

GN Neurospora crassa.

OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

CC Sordariales; Sordariaceae; Neurospora.

CC NCBI\_TaxID=5141;

OX NCBI\_TaxID=5141;

RN [1]

RP MEDLINE FROM N.A.

RX MEDLINE-93101148; PubMed-1361212;

RA Heintz K., Palme K., Diefenthal T., Russo V.E.A.;

RT "The Nyepl1 gene from Neurospora crassa is located on chromosome 2:

RT molecular cloning and structural analysis.";

RL Mol. Gen. Genet. 235:413-421(1992).

CC [2]

RP SEQUENCE FROM N.A.

RC STRAIN-74-OR23-1A;

RA Schulte U., Aign V., Hohnsels J., Brandt P., Farthmann B., Holland R.,



RA Miyake G., Mewes H.-W., Mannhaupt G.;  
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR  
 CC TRAFFIC (BY SIMILARITY).  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: S1252; AAB24564.1;  
 CC EMBL: AL356324; CAB92031.1;  
 CC PIR: S30096; S30096.  
 CC HSSP: P01112; 1PLT.  
 CC InterPro: IPR003579; GTPase\_Rab.  
 CC InterPro: IPR001806; Ras\_trnsfrmng.  
 CC InterPro: IPR005225; Small\_GTP.  
 CC Pfam: PF000071; Ras\_1.  
 CC PRINTS: PR00449; RASRNSFRNG.  
 CC SMART: SM00175; RAB; 1.  
 CC TIGRfam: TIGR00231; small\_gtp; 1.  
 CC GTP-binding; Lipoprotein; Prenylation; Protein transport.  
 CC KW NP\_BIND 15 22 GTP (BY SIMILARITY).  
 CC FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 CC FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 CC FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).  
 CC FT LIPID 202 202 GERANYL-GERANYL (BY SIMILARITY).  
 CC FT LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).  
 CC SQ SEQUENCE 203 AA; 22476 MW; 20F51AB3400EFID CRC64;  
 Query Match 74.3%; Score 658.5; DB 1; Length 203;  
 Best Local Similarity 66.5%; Pred. No. 5.3e-48;  
 Matches 135; Conservative 11; Mismatches 24; Indels 33; Gaps 2;  
 QY 4 MNPEYDYELFKLLIGDSCVGRKSCLLRFADDTYESYSTIGVDKIRTIENLGKTIKQ 63  
 DB 1 MNPEYDYELFKLLIGDSCVGRKSCLLRFADDTYESYSTIGVDKIRTIENLGKTIKQ 60  
 QY 64 I-----ESFNNVKOMLOEIDRYASENVKLLVG 91  
 DB 61 IMPYAGQERFRTITSSYRGAHGICVYDVTDMDSFNNVKOMLOEIDRYASENVKLLVG 120  
 QY 92 NKCDLTKRKVVDYTTAKAFADSLGIPLETSAKNATNVEGSPWMAAEIKKRMGPGATAG 151  
 DB 121 NKSDMTSKKVEYEVAKAFADSLGIPLETSAKNATNVEGAFETMARQIKRMGSGSTAIN 180  
 QY 152 GAERK-NVYKQSTPYKOSGGCC 173  
 DB 181 NTKASVAVSPGHGVSNNSSGCC 203  
 RESULT 7  
 YPFL\_CHLRE STANDARD; PRT; 203 AA.  
 ID YPFL\_CHLRE  
 AC Q39571;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE GTP-binding protein YPFLC1.  
 GN YPFLC1.  
 OS Chlamydomonas reinhardtii.  
 CC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 CC Chlamydomonadaceae; Chlamydomonas.  
 CC NCB1\_TaxID=3055;  
 CC NCB1\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CW15;  
 RX MEDLINE=95309723; PubMed=7789809;  
 RA Dietmaier W., Fabry S., Huber H., Schmitt R.;

RT "Analysis of a family of ypt genes and their products from  
 RT Chlamydomonas reinhardtii.";  
 RL Gene 158:41-50(1995).  
 CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR  
 CC TRAFFIC (BY SIMILARITY).  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U13168; AAB2727.1;  
 CC HSSP: P05713; 3RAB.  
 CC InterPro: IPR003579; GTPase\_Rab.  
 CC InterPro: IPR001806; Ras\_trnsfrmng.  
 CC InterPro: IPR005225; Small\_GTP.  
 CC Pfam: PF000071; Ras\_1.  
 CC PRINTS: PR00449; RASRNSFRNG.  
 CC SMART: SM00175; RAB; 1.  
 CC TIGRfam: TIGR00231; small\_gtp; 1.  
 CC GTP-binding; Lipoprotein; Prenylation; Protein transport;  
 CC KW Multigene family.  
 CC FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 CC FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 CC FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 CC FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).  
 CC FT LIPID 202 202 GERANYL-GERANYL (BY SIMILARITY).  
 CC FT LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).  
 CC SQ SEQUENCE 203 AA; 22599 MW; DD22420FE2F35EID CRC64;  
 Query Match 73.0%; Score 646.5; DB 1; Length 203;  
 Best Local Similarity 65.7%; Pred. No. 5.3e-47;  
 Matches 134; Conservative 9; Mismatches 26; Indels 35; Gaps 3;  
 QY 4 MNPEYDYELFKLLIGDSCVGRKSCLLRFADDTYESYSTIGVDKIRTIENLGKTIKQ 63  
 DB 1 MNPEYDYELFKLLIGDSCVGRKSCLLRFADDTYESYSTIGVDKIRTIENLGKTIKQ 60  
 QY 64 I-----ESFNNVKOMLOEIDRYASENVKLLVG 91  
 DB 61 IMPYAGQERFRTITSSYRGAHGICVYDVTDMDSFNNVKOMLOEIDRYASENVKLLVG 120  
 QY 92 NKCDLTKRKVVDYTTAKAFADSLGIPLETSAKNATNVEGSPWMAAEIKKRMGPGATAG 149  
 DB 121 NKSDMTSKKVEYEVAKAFADSLGIPLETSAKNATNVEGAFETMARQIKRMGSGSTAIN 180  
 QY 150 AGAERKSNVYKQSTPYKOSGGCC 173  
 DB 181 KAGGPVVRPO-BGRPINSKSSGCC 203  
 RESULT 8  
 YPFL\_SCHPO STANDARD; PRT; 203 AA.  
 ID YPFL\_SCHPO  
 AC P11620;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ras-related protein YPFL.  
 GN YPFL OR SPBC1703.10.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomyces.  
 CC NCB1\_TaxID=4896;  
 CC NCB1\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=90228338; PubMed=2328721;  
 RX Miyake S., Yamamoto M.;

RT "Identification of ras-related, YPT family genes in  
RT Schizosaccharomyces pombe".  
RL EMBL J. 9:1417-1422(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-89296485; PubMed-2740223;  
RA Fawell E., Hook S., Armstrong J.;  
RT "Nucleotide sequence of a gene encoding a YPT1-related protein from  
RT Schizosaccharomyces pombe".  
RL Nucleic Acids Res. 17:4373-4373(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RX MEDLINE-21848401; PubMed-11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Frazer A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leacher S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,  
RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Meisel D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Mambitt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garçon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe".  
RL Nature 415:871-880(2002).  
RN [4]  
RP ISOPRENOID.  
RX MEDLINE-92283842; PubMed-1597466;  
RA Newman C.M., Giannakouras T., Hancock J.F., Fawell E.H., Armstrong J.,  
RT Magee A.L.;  
RT "Post-translational processing of Schizosaccharomyces pombe YPT  
RT proteins".  
RL J. Biol. Chem. 267:11329-11336(1992).  
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR  
CC TRAFFIC (BY SIMILARITY).  
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X52099; CA36319.1; -;  
DR EMBL: X15082; CA33192.1; -;  
DR EMBL: AL136536; CAB66454.1; -;  
DR PIR: S04590; S04590.  
DR PIR: S10025; S10025.  
DR HSSP: P05713; 3RAB.  
DR InterPro: IPR003579; GTPase\_Rab.  
DR InterPro: IPR001806; Ras\_trnsfmg.  
DR InterPro: IPR005225; Small\_GTP.  
DR Pfam: PF00071; ras; 1.  
DR PRINTS: PR00449; RASTRNSFRMG.  
DR SMART: SM00175; RAB; 1.  
DR TIGRfams: TIGR00231; small\_GTP; 1.

KW GTP-binding; Lipoprotein; Prenylation; Protein transport.  
FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).  
FT LIPID 202 202 GERANYL-GERANYL.  
FT LIPID 203 203 GERANYL-GERANYL.  
FT CONFLICT 1 4 MNPE -> MNANINR (IN REF. 2).  
SQ SEQUENCE 203 AA; 22816 MW; 59B3524F42B3094 CRC64;  
  
Query Match 73.0%; Score 646.5; DB 1; Length 203;  
Best Local Similarity 64.5%; Pred. No. 5.3e-47;  
Matches 131; Conservative 16; Mismatches 23; Indels 33; Gaps 2;  
  
OY 4 MNPEVDYLFKLLIDSGKSCLLRPADPDYTESYSTIGVDPRKRTIEIDGKTIKQ 63  
DB 1 MNPEVDYLFKLLIDSGKSCLLRPADPDYTESYSTIGVDPRKRTIEIDGKTIKQ 60  
OY 64 I-----ESFNWYKWLQIDRYASENVKLLVG 91  
DB 61 IMDTAGQERFRITSSYYRGAGHIIYDVTDQDSFNWYKWLQIDRYAVGVNRLVG 120  
OY 92 NKCDLTKKYVDYTAKEPADSLGPELETSKKNATNQSMTMAAETKRMGCATAG 151  
DB 121 NKSDVDKVKVEYSYAKEPADSLNIPLETSAKDSINWQAFIWSROIKEKMGNTPAS 180  
OY 152 GAESKNVXI-OSTPVKSGGCGC 173  
DB 161 SNAKSVAVGQGTNNVSQSSNCC 203  
  
RESULT 9  
ID YPT1\_VOLCA STANDARD; PRT; 203 AA.  
AC P31584;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE GTP-binding protein ypt1.  
GN YPT1.  
OS Volvox carterl.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Volvocaceae; Volvox.  
OX NCBI\_TaxID=3067;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-f. Nagariensis / HK10;  
RX MEDLINE-92380499; PubMed-1511889;  
RA Fabry S., Nass N., Huber H., Palme K., Jaenicke L., Schmitt R.;  
RT "The ypt1 gene encodes a small G-protein in the green alga Volvox  
RT carterl: gene structure and properties of the gene product".  
RL Gene 118:153-162(1992).  
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR  
CC TRAFFIC (BY SIMILARITY).  
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M93438; AAA34255.1; -;  
DR PIR: JC1247; JC1247.  
DR HSSP: P05713; 3RAB.  
DR InterPro: IPR003579; GTPase\_Rab.  
DR InterPro: IPR001806; Ras\_trnsfmg.  
DR InterPro: IPR005225; Small\_GTP.  
DR Pfam: PF00071; ras; 1.  
DR PRINTS: PR00449; RASTRNSFRMG.  
DR SMART: SM00175; RAB; 1.

DR	TIGR03111	small GTP, 1.	Protein transport;
DR	GTP-binding; Lipoprotein; Prenylation;		
DR	Multi-gene family.		
DR	NP_BIND	15	22
DR	NP_BIND	63	67
DR	NP_BIND	121	124
DR	DOMAIN	37	45
DR	LIPID	202	202
DR	LIPID	203	203
DR	SEQUENCE	203 AA; 225532 MM; CDL43BADB92BAC65 CRC64;	
DR	Query Match	72.5%; Score 642.5; DB: 1; Length 203;	
DR	Best Local Similarity	65.0%; Pred. No. 1.1e-46;	
DR	Matches 134; Conservative	6; Mismatches 27; Indels 39; Gaps 3;	
DR	Y	4	MNEVDYLFKLLIGDSGVKSCLLRFADDTYTESYITIGVDFKRTIELDQRTIKLQ 63
DR	Db	1	MNEVDYLFKLLIGDSGVKSCLLRFADDTYTESYITIGVDFKRTIELDQRTIKLQ 60
DR	Y	64	I-----ESFNNWQMLQEDRYASENNKLLVG 91
DR	Db	61	IMDTAGOEFTTISSYRGAGHGIIVYDWDSEFNNWQMLQEDRYASENNKLLVG 120
DR	Y	92	NKCDLFTKRVVDYTFKREFASLGIPLFETSANNTNEOSFMTMAAEIKRMG----CG 147
DR	Db	121	NKSDLCKKRVVDYQAKAKADAEICGPLFETSANNTNEQAFMTMAAEIKRMASGVPVP 180
DR	Y	148	ATAGSAKSNYKISQIPVQSGGGCC 173
DR	Db	181	KPGGVPVPRF---EGRFINNKSSSC 203
DR	RESULT 10		
DR	YPTL_PHYIN		
DR	ID	YPTL_PHYIN	STANDARD; PRT: 201 AA.
DR	AC	001890;	
DR	DT	01-NOV-1997 (Rel. 35; Created)	
DR	DT	01-NOV-1997 (Rel. 35; Last sequence update)	
DR	DT	01-NOV-1997 (Rel. 35; Last annotation update)	
DR	DE	Ras-like GTP-binding protein 1P71.	
DR	GN	Phytophthora infestans (Potato late blight fungus).	
DR	OC	Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;	
DR	OC	Phytophthora.	
DR	OX	NCBI_TaxID=4787;	
DR	RN	(1)	
DR	RC	SEQUENCE FROM N.A.	
DR	RC	STRAIN=ATCC 52009;	
DR	RX	MEDLINE=97128773; PubMed=8973313;	
DR	RA	Chen Y., Rosdy R.;	
DR	RT	Characterization of a phytophthora infestans gene involved in vesicle transport.;	
DR	RL	Gene 181:89-94(1996).	
DR	CC	-1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR TRAFFIC FROM ER TO GOLGI (BY SIMILARITY).	
DR	CC	-1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAS SUBFAMILY.	
DR	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).	
DR	CC	EMBL: U04074; AAA0355.1; -	
DR	DR	HSSP: P05713; RAB	
DR	DR	InterPro: IPR003579; GTPase_Rab.	
DR	DR	InterPro: IPR001806; Ras_Ltrnsfmg.	
DR	DR	InterPro: IPR005225; Small_GTP.	
DR	DR	Pfam: PF00071; ras; 1.	
DR	DR	PRINTS: PR00449; RASTRNSPRMG.	
DR	DR	SMART: SM00175; RAB; 1.	

DR TRIGERMS; TRIGR00231; small GTP; 1.  
 KW GTP-binding; Lipoprotein; Preonitiation; Protein transport.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 PROTECTOR REGION (PROBABLE).  
 FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).  
 FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).  
 FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).  
 SQ SEQUENCE 201 AA; 22292 MW; 89E553A693C8BAC5 CRC64;  
  
 Query Match 71.3%; Score 631.5; DB 1; Length 201;  
 Best Local Similarity 67.5%; Pred. No. 9.2e-46;  
 Matches 139; Conservative 8; Mismatches 18; Indels 41; Gaps 5  
  
 QY 4 MNPEYDLERKLLIGDSGVKSCOLLIRPADYTESYSTIGVDKFKRTIELDGKTIQLQ 63  
 Db 1 MNPEYDLERKLLIGDSGVKSCOLLIRPADYTESYSTIGVDKFKRTIELDGKTIQLQ 60  
 QY 64 I-----ESFNNNKMKQLIEDIRASENVKLLLVG 91  
 Db 1 IMPTAQGEFFRTTSSYYRGAGIIIVVDYDSEFNNNKMKQLIEDIRASENVKLLLVG 120  
 QY 92 NKCDLTRKRVDDYTTAAEFADSGIPIETSAKNANNVQSPFMMAEKRMKGCATG 151  
 Db 121 NKSDLTRKRVDDYTTAAEFADSGIPIETSAKNANNVQSPFMMAEKRMKGCATG 177  
 QY 152 GAESNVKTIQSPYKQ--SGSG--CC 173  
 Db 178 VAPRAGVRL--TPGQVPSNGSGKCC 201  
  
 RESULT 11  
 ARAS\_AARATH STANDARD: PRT; 256 AA.  
 ID ARAS\_AARATH STANDARD: PRT; 256 AA.  
 DC 628186;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 15-JUN-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ras-related protein ARAS-5.  
 GN AAA-5 OR ATIG02130 OR T7133.6.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN 1  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altif F.H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Chuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.B., Creasy T.H., Dewar K.,  
 RA Dunn P., Etlip P., Feldblum T.V., Feng J.-D., Feng B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huttar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremetska I., Kuritz D.B., Kuan A., Lam B.,  
 RA Langlois-Hopson S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-F.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros S., Maletti R., Marzilli A.,  
 RA Miltsecher T., Miranda M., Nguyen M., Newman W.C., Osborne B.I.,  
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Sherman W.C., Soltner B.I.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Taiton L.J., Tambung J., Town M.J., Town C.D.,  
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana.  
 RL Nature 408:816-820(2000).  
 (2)  
 SEQUENCE OF 64-258 FROM N.A.  
 RC STRAIN=Various strains; TISSUE=Leaf;  
 RX MEDLINE=92084144; PubMed=1748311;  
 RA Anal T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,

RA Matsui M.;  
 RT "Isolation and analysis of cDNAs encoding small GTP-binding proteins  
 of Arabidopsis thaliana."  
 RL Gene 108:259-264(1991).  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U89959; AAC24370.1; -  
 CC EMBL: D01027; BAA00832.1; -  
 CC PIR: PS0279; PS0279.  
 CC HSSP: P05713; 3RAB.  
 CC InterPro: IPR003579; GTPase\_Rab.  
 CC InterPro: IPR001806; Ras\_tnsftrng.  
 CC InterPro: IPR005225; Small\_GTP.  
 CC Pfam: PF00071; Ras; 1.  
 CC PRINTS: PRO0449; RASTRNSFRNG.  
 CC SMART: SM00175; RAB; 1.  
 CC TIGRFAMs: TIGR00231; small\_GTP; 1.  
 CC GTP-binding; Lipoprotein; Prenylation; Multigene family.  
 KW NP\_BIND 70 77  
 FT NP\_BIND 118 122 GTP (BY SIMILARITY).  
 FT NP\_BIND 176 179 GTP (BY SIMILARITY).  
 FT DOMAIN 92 100 EFFECTOR REGION (BY SIMILARITY).  
 FT LIPID 255 255 GERANYL-GERANYL (BY SIMILARITY).  
 FT LIPID 256 256 GERANYL-GERANYL (BY SIMILARITY).  
 SQ SEQUENCE 258 AA; 29169 MW; 34D278A0C504638P CRC64;  
 Query Match 69.9%; Score 619.5; DB 1; Length 258;  
 Best Local Similarity 61.6%; Pred. No. 1.2e-44;  
 Matches 125; Conservative 15; Mismatches 30; Indels 33; Gaps 2;  
 QY 3 SMNPEDYLFKLLIGDSGVGKSCLLRFADDTYESTYSTIGVDKIRITELDKITKL 62  
 DB 55 TMNPEYDYLKLLIGDSGVGKSCLLRFSDSYESTYSTIGVDKIRITVEQDKITKL 114  
 QY 63 QI-----ESFNNVQWLOEIDRYASENVNKLIV 90  
 DB 115 QIMDTAGOEFRRTTSSYYRGAGHIIIVDYDTDESFPNNVQWLOEIDRYASDNVKNLIV 174  
 QY 91 GNKCDLTKKVVDTTAKFADSLGIPLETSAKNATVQESFMTMAEIKRKGPGATA 150  
 DB 175 GNKCDLTKKVVDTTAKFADSLGIPLETSAKNATVQESFMTMAEIKRKGPGATA 150  
 QY 151 GGAESNVKIQSTPVKQSGGCC 173  
 DB 235 NNARPTVQIRGQPVNOK-NGCC 256  
 RESULT 12  
 ID YPT2\_MAIZE STANDARD; PRT; 203 AA.  
 AC Q05737;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE GTP-binding protein YPT2.  
 GN YPT2.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Palicoudeae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Coleoptile;  
 RX MEDLINE=92115746; PubMed=1731354;

RA Palme K., Diefenthal T., Vingron M., Sander C., Schell J.;  
 RT "Molecular cloning and structural analysis of genes from Zea mays  
 (L.) coding for members of the ras-related ypt gene family."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:787-791(1992).  
 CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR  
 CC TRAFFIC (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: ITS EXPRESSION IS WEAK IN STEMS, HIGHER IN  
 CC ROOTS, LEAVES AND COLEOPTILES, BUT HIGHEST IN FLOWERS.  
 CC -1- PFM: THE PALMITIC ACID IS REQUIRED FOR MEMBRANE ATTACHMENT AND  
 CC BIOLOGICAL FUNCTION.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X63278; CAA44919.1; -  
 CC PIR: B38202; B38202.  
 CC HSSP: P05713; 3RAB.  
 CC MaizeDB: 78605; -  
 CC InterPro: IPR003579; GTPase\_Rab.  
 CC InterPro: IPR001806; Ras\_tnsftrng.  
 CC InterPro: IPR005225; Small\_GTP.  
 CC Pfam: PF00071; Ras; 1.  
 CC PRINTS: PRO0449; RASTRNSFRNG.  
 CC SMART: SM00175; RAB; 1.  
 CC TIGRFAMs: TIGR00231; small\_GTP; 1.  
 CC GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport.  
 KW NP\_BIND 15 22  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).  
 FT LIPID 200 200 PALMITATE (BY SIMILARITY).  
 FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).  
 SQ SEQUENCE 203 AA; 22475 MW; E2413267ACD1B8A CRC64;  
 Query Match 69.8%; Score 618.5; DB 1; Length 203;  
 Best Local Similarity 63.4%; Pred. No. 1.1e-44;  
 Matches 128; Conservative 11; Mismatches 30; Indels 33; Gaps 2;  
 QY 4 MNPEYDYLKLLIGDSGVGKSCLLRFADDTYESTYSTIGVDKIRITELDKITKLQ 63  
 DB 1 MNPEYDYLKLLIGDSGVGKSCLLRFADDTYESTYSTIGVDKIRITVEQDKITKLQ 60  
 QY 64 I-----ESFNNVQWLOEIDRYASENVNKLIV 91  
 DB 61 IMDTAGOEFRRTTSSYYRGAGHIIIVDYDTDESFPNNVQWLOEIDRYASDNVKNLIV 120  
 QY 92 NKCDLTKKVVDTTAKFADSLGIPLETSAKNATVQESFMTMAEIKRKGPGATA 151  
 DB 121 NKCDLTKKVVDTTAKFADSLGIPLETSAKNATVQESFMTMAEIKRKGPGATA 180  
 QY 152 GAESNVKIQSTPVKQSGGCC 173  
 DB 181 NARAPTQIRGQPVNOK-TSCC 201  
 RESULT 13  
 ID RIC1\_ORYSA STANDARD; PRT; 202 AA.  
 AC P40392;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ras-related protein RIC1.  
 GN RIC1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=84093563; PubMed=6318115;
RA Gallwitz D., Donath C., Sander C.;
RT "A yeast gene encoding a protein homologous to the human c-ha/bas
RL Nature 306:704-707(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5286C; AB972;
RA MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT Analysis of the nucleotide sequence of chromosome VI from
RL Saccharomyces cerevisiae.
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP PALMITOYLATION.
RA MEDLINE=88296440; PubMed=3042385;
RA Molenaar C.M.T., Prange R., Gallwitz D.;
RT "A carboxyl-terminal cysteine residue is required for palmitic acid
RT binding and biological activity of the ras-related yeast Ypt1
RT protein."
RL EMOB J. 7:971-976(1988).
RN [4]
RP MUTAGENESIS, AND POSSIBLE FUNCTION.
RA MEDLINE=88223356; PubMed=3286011;
RA Schmitt H.D., Puzicha M., Gallwitz D.;
RT "Study of a temperature-sensitive mutant of the ras-related Ypt1 gene
RT product in yeast suggests a role in the regulation of intracellular
RT calcium."
RL Cell 53:635-647(1988).
RN [5]
RP MUTAGENESIS OF EFFECTOR REGION.
RA MEDLINE=91184123; PubMed=2009858;
RA Becker J., Tan T., Trepte H.-R., Gallwitz D.;
RT "Mutational analysis of the putative effector domain of the
RT GTP-binding Ypt1 protein in yeast suggests specific regulation by a
RT novel GAP activator."
RL EMOB J. 10:785-793(1991).
RN [6]
RP -1- FUNCTION: YPT1 IS ESSENTIAL FOR CELL VIABILITY. THE LOSS OF YPT1
RN RESULTS IN CYTOSKELETAL AND MITOTIC LESIONS. YPT1 IS REQUIRED FOR
RN TRANSIT OF PROTEINS FROM THE ER THROUGH GOLGI COMPARTMENT.
RN -1- THE PALMITIC ACID IS REQUIRED FOR MEMBRANE ATTACHMENT AND
RN BIOLOGICAL FUNCTION.
RN -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN
RN FRBS Left. 332:282-286(1993).
RN -1- FUNCTION: POSSESSES GTPASE ACTIVITY.
RN -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
RN
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN
RN EMBL: S66160; AAB28535.1; -.
RN PIR: S38740; S38740.
RN HSSP: P01112; 1PLK.
RN InterPro: IPR003579; GTPase_Rab.
RN InterPro: IPR001806; Ras_transfmg.
RN InterPro: IPR005225; Small_GTP.
RN Pfam: PF00071; Ras; 1.
RN PRINTS: PR00449; RASTRNSFRNG.
RN SMART: SM00175; RAB; 1.
RN TIGRFAm: TIGR00231; small_GTP; 1.
RN GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport.
RN NP_BIND 15 22 GTP (BY SIMILARITY).
RN NP_BIND 63 67 GTP (BY SIMILARITY).
RN NP_BIND 121 124 GTP (BY SIMILARITY).
RN DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
RN LIPID 200 200 PALMITATE (BY SIMILARITY).
RN LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
RN SEQUENCE 202 AA; 22561 MW; 5A1CB5F3256B659 CRC64;
RN
RN Query Match 69.2%; Score 613.5; DB 1; Length 202;
RN Best Local Similarity 61.9%; Pred. No 2; 9e-44;
RN Matches 125; Conservative 12; Mismatches 32; Indels 33; Gaps 2;
RN
RN 4 MPEPDIKRLILIGSGVSKLLRFADDTYESISTGVDFKRIELDKTITLQ 63
RN 1 MPEPDIKRLILIGSGVSKLLRFADDTYESISTGVDFKRIELDKTITLQ 60
RN 64 I-----ESNNVKOMLOEIDRVSNNKLLVG 91
RN 61 IWDNAGQERRTTSSYRGAGIIVYDVTDQESNNVKOMLEIDRVASNNKLLVG 120
RN 92 NKCDLTKKKVVDYTAKEFADSLGIPLETSAKNTNVEOSFMTMAEIKRMGGATAG 151
RN 121 NKCDLARNRVSYEAKLALDEIGIPLETSAKNTNVEKAFMTMAGEIKRMGGATAG 180
RN 152 GAESNVKIQSTPVGKSGGCC 173
RN 181 ASKPATVQMPROPVPAQO--SSCC 201
RN
RN RESULT 14
RN YPT1_YEAST STANDARD; PRT; 206 AA.
RN AC P01123;
RN DT 21-JUL-1986 (Rel. 01, Created)
RN DT 01-NOV-1995 (Rel. 32, Last sequence update)
RN DT 15-JUN-2002 (Rel. 41, Last annotation update)
RN DE GTP-binding protein YPT1 (Protein YP2).
RN GN YPT1 OR YP2 OR YF1038C.
RN OS Saccharomyces cerevisiae (Baker's yeast).

```

```

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=84093563; PubMed=6318115;
RA Gallwitz D., Donath C., Sander C.;
RT "A yeast gene encoding a protein homologous to the human c-ha/bas
RL Nature 306:704-707(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5286C; AB972;
RA MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT Analysis of the nucleotide sequence of chromosome VI from
RL Saccharomyces cerevisiae.
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP PALMITOYLATION.
RA MEDLINE=88296440; PubMed=3042385;
RA Molenaar C.M.T., Prange R., Gallwitz D.;
RT "A carboxyl-terminal cysteine residue is required for palmitic acid
RT binding and biological activity of the ras-related yeast Ypt1
RT protein."
RL EMOB J. 7:971-976(1988).
RN [4]
RP MUTAGENESIS, AND POSSIBLE FUNCTION.
RA MEDLINE=88223356; PubMed=3286011;
RA Schmitt H.D., Puzicha M., Gallwitz D.;
RT "Study of a temperature-sensitive mutant of the ras-related Ypt1 gene
RT product in yeast suggests a role in the regulation of intracellular
RT calcium."
RL Cell 53:635-647(1988).
RN [5]
RP MUTAGENESIS OF EFFECTOR REGION.
RA MEDLINE=91184123; PubMed=2009858;
RA Becker J., Tan T., Trepte H.-R., Gallwitz D.;
RT "Mutational analysis of the putative effector domain of the
RT GTP-binding Ypt1 protein in yeast suggests specific regulation by a
RT novel GAP activator."
RL EMOB J. 10:785-793(1991).
RN [6]
RP -1- FUNCTION: YPT1 IS ESSENTIAL FOR CELL VIABILITY. THE LOSS OF YPT1
RN RESULTS IN CYTOSKELETAL AND MITOTIC LESIONS. YPT1 IS REQUIRED FOR
RN TRANSIT OF PROTEINS FROM THE ER THROUGH GOLGI COMPARTMENT.
RN -1- THE PALMITIC ACID IS REQUIRED FOR MEMBRANE ATTACHMENT AND
RN BIOLOGICAL FUNCTION.
RN -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN
RN EMBL: X00209; CAA25036.1; -.
RN EMBL: D50617; BAA09201.1; -.
RN PIR: A01374; TVYQ2.
RN HSSP: P05713; 3RAB.
RN SGD: S0001856; YPT1.
RN InterPro: IPR003579; GTPase_Rab.
RN InterPro: IPR001806; Ras_transfmg.
RN InterPro: IPR005225; Small_GTP.
RN Pfam: PF00071; Ras; 1.
RN PRINTS: PR00449; RASTRNSFRNG.
RN SMART: SM00175; RAB; 1.
RN TIGRFAm: TIGR00231; small_GTP; 1.
RN GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport; Golgi stack;
RN Endoplasmic reticulum; Palmitate.

```

```

FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT LIPID 205 205 PALMITATE.
FT LIPID 206 206 GERANYL-GERANYL.
FT MUTAGEN 37 37 Y->F: NO CHANGE.
FT MUTAGEN 39 39 S->A: NO CHANGE.
FT MUTAGEN 40 40 T->S: NO CHANGE.
FT MUTAGEN 41 41 I->M: LETHAL.
FT MUTAGEN 43 43 V->E: NO CHANGE.
FT MUTAGEN 44 44 D->N: TEMPERATURE-SENSITIVE PHENOTYPE.
FT MUTAGEN 121 121 N->I: LETHAL.
FT CONFILIT 171 171 E->Q (IN REF. 1).
SQ SEQUENCE 206 AA; 23214 MW; F8C704F6BF2D227B CRC64;

Query Match 64.4%; Score 571; DB 1; Length 206;
Best Local Similarity 56.3%; Pred. No. 1e-40;
Matches 116; Conservative 21; Mismatches 33; Indels 36; Gaps 2;

QY 4 MNPEYDYLFLKLLIDSGVSKSCLLRADPTYSYSTIGVDFKIRTIELDKTKTLQ 63
DB 1 MNPEYDYLFLKLLIDSGVSKSCLLRADPTYSYSTIGVDFKIRTIELDKTKTLQ 60
QY 64 I-----ESFNNVQWLOEIDRYASENVKLLVG 91
DB 61 IWDTAGQERFRTTSSYYRGSHGIIIVYDVTDOESFNVKMWLOEIDRYASTVTKLLVG 120
QY 92 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGP-----G 147
DB 121 NKCDLTKDRVVEYDPAKFAADANKMPLETSLDSTINVEDAFLTARQIKESMSQONLNE 180
QY 148 AFAGAEKSNVKIQSTPVKQSGGCC 173
DB 181 TTQKKEDKGNVNLKGOSLTNTGGGCC 206

RESULT 15
RBLA_DICDI STANDARD; PRT; 167 AA.
AC P34139;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Ras-related protein Rab1A.
GN RAB1A OR RAB1A.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=94124041; PubMed=8294041;
RA Bush J.M. IV, Franek K., Daniel J.M., Spiegelman G.B., Weeks G.,
RT Cardelli J.A.;
RT "Cloning and characterization of five novel Dictyostelium discoideum
RT rad-related genes.";
RL Gene 136:55-60(1993).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; L21009; AAC37385.1; -.
CC HSSP; P05713; 3RAB.
CC DictyDb; DD05055; raba.
CC InterPro; IPR003579; GTPase_Rab.
CC InterPro; IPR001806; Ras_transfmg.

```

```

DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_GTP; 1.
KW GTP-binding.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
SQ SEQUENCE 167 AA; 18943 MW; 72480B04D943BD8D CRC64;

Query Match 61.2%; Score 542; DB 1; Length 167;
Best Local Similarity 67.7%; Pred. No. 2e-38;
Matches 111; Conservative 10; Mismatches 11; Indels 32; Gaps 1;

QY 4 MNPEYDYLFLKLLIDSGVSKSCLLRADPTYSYSTIGVDFKIRTIELDKTKTLQ 63
DB 1 MNPEYDYLFLKLLIDSGVSKSCLLRADPTYSYSTIGVDFKIRTIELDKTKTLQ 60
QY 64 I-----ESFNNVQWLOEIDRYASENVKLLVG 91
DB 61 IWDTAGQERFRTTSSYYRGAGHIIIVYDVTDKLTFENVRQWLOEIDRFACENVKLLVG 120
QY 92 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVEOSFMT 135
DB 121 NKSDLVARKVVDENTAKAFADSLQIPLETSAKOSTNVEQAFMT 164

```

Search completed: June 18, 2003, 15:21:38  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:18:29 ; Search time 80 Seconds

(without alignments)  
445,577 Million cell updates/sec

Title: US-09-820-003b-2

Perfect score: 1 MSSMNPEDYLFKLLIGDS.....EKSNNKIQSTPYKSGSGCC 173

Sequence: 1 MSSMNPEDYLFKLLIGDS.....EKSNNKIQSTPYKSGSGCC 173

Scoring table: BLOSUM62

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_ORNITHINE:\*  
8: SP\_ORNITHINE:\*  
9: SP\_ORNITHINE:\*  
10: SP\_ORNITHINE:\*  
11: SP\_ORNITHINE:\*  
12: SP\_ORNITHINE:\*  
13: SP\_ORNITHINE:\*  
14: SP\_ORNITHINE:\*  
15: SP\_ORNITHINE:\*  
16: SP\_ORNITHINE:\*  
17: SP\_ORNITHINE:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	756.5	86.5	201	4 Q9H0U4	Q9H0U4 homo sapien
2	759.5	85.7	201	11 Q9H0G1	Q9H0G1 mus musculu
3	700	79.0	141	4 Q9H0G1	Q9H0G1 homo sapien
4	693	77.1	202	5 Q18332	Q18332 drosophila
5	671	75.1	202	5 Q18332	Q18332 drosophila
6	660	74.5	203	3 Q9H0A6	Q9H0A6 caenorhabdi
7	653.3	74.0	201	3 Q9H0A6	Q9H0A6 caenorhabdi
8	653.3	73.8	201	3 Q9H0A6	Q9H0A6 caenorhabdi
9	644	72.5	202	10 Q40203	Q40203 lotus japon
10	640	72.5	202	10 Q40203	Q40203 lotus japon
11	636.5	71.8	202	3 Q9H0A6	Q9H0A6 caenorhabdi
12	634.5	71.6	202	10 Q9H0A6	Q9H0A6 caenorhabdi
13	633	71.4	202	10 Q9H0A6	Q9H0A6 caenorhabdi
14	632.5	71.4	203	10 Q24112	Q24112 nicotiana p
15	632	71.3	202	10 Q9SXT5	Q9SXT5 cicer ariet
16	632	71.3	202	10 Q9SXT5	Q9SXT5 cicer ariet

17	628.5	70.9	203	10 Q40204	Q40204 lotus japon
18	627.5	70.8	203	10 Q40569	Q40569 nicotiana t
19	624.5	70.5	203	10 Q40205	Q40205 lotus japon
20	624	70.4	129	4 Q9Y372	Q9Y372 homo sapien
21	623	70.3	202	10 Q39845	Q39845 glycine max
22	622.5	70.3	203	10 Q9M7P5	Q9M7P5 capsicum an
23	622.5	70.3	203	10 Q41338	Q41338 lycopersico
24	622.5	70.3	203	10 Q41340	Q41340 lycopersico
25	621.5	70.1	203	10 Q949E2	Q949E2 oryza sativ
26	619	69.9	202	10 Q08155	Q08155 pisum sativ
27	618.5	69.8	203	10 Q940Z7	Q940Z7 arabidopsis
28	616	69.5	202	10 Q08153	Q08153 pisum sativ
29	612	69.1	221	10 Q23594	Q23594 arabidopsis
30	608	68.6	203	10 Q8RU63	Q8RU63 oryza sativ
31	598	67.5	218	10 Q39861	Q39861 glycine max
32	596.5	67.3	207	3 Q96W97	Q96W97 candida alb
33	595.5	66.6	202	10 Q92RH6	Q92RH6 candida alb
34	589	66.5	201	10 Q08152	Q08152 pisum sativ
35	589	66.5	451	10 Q91G53	Q91G53 oryza sativ
36	587	66.3	201	10 Q43462	Q43462 glycine max
37	585	65.9	202	10 Q41339	Q41339 lycopersico
38	581	65.6	201	10 Q40201	Q40201 lotus japon
39	580	65.5	200	5 Q25324	Q25324 lotus japon
40	574	64.8	206	10 Q38274	Q38274 leishmania
41	569.5	64.3	205	10 Q38274	Q38274 leishmania
42	565	64.2	194	10 Q40202	Q40202 lotus japon
43	563	63.3	206	3 Q9H0M5	Q9H0M5 kluveromyc
44	503	56.8	200	5 Q9BHNI	Q9BHNI plasmodium
45	484	54.6	205	5 Q9BLF3	Q9BLF3 entamoeba h

## ALIGNMENTS

RESULT 1  
ID Q9H0U4 PRELIMINARY: PRT; 201 AA.  
AC Q9H0U4  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 22.2 kDa protein.  
GN DKFZ5641172.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN.  
RX MEDLINE=2154917; PubMed=11230166;  
RA Wieman S., Wall B., Wellenreuther R., Gassenhuber J., Glasl S.,  
RA Ansgore W., Boescher M., Bloeker H., Bauersachs S., Blum H.,  
RA Rauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
RA Meves H.W., Oetemaelder B., Obermaier B., Tampe J., Heuner D.,  
RA "Towards a Catalog of Human Genes and Proteins: Sequencing and  
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";  
RL Genome Res. 11:422-435(2001).  
CC 1- SIMILARITY: TO RAB PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
EMBL: AL13635; CAB6570.1; -;  
HSP: P05713; RAB.  
DR InterPro: IPR003579; GTPase\_Rab.  
DR InterPro: IPR001806; Ras\_Trasfmrng.  
DR InterPro: IPR002078; Sig54\_Interact.  
DR InterPro: IPR005225; Small\_GTP.  
DR Pfam: PF00071; Ras; 1.  
DR PRINTS: PR00449; RASTRNSFRNG.  
DR SMART: SM00175; RAB; 1.  
DR TIGRFAMs: TIGR00231; small\_GTP; 1.  
DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
KW GTP-binding; Hypothetical protein; Lipoprotein.  
SQ SEQUENCE 201 AA: 22171 MW: 9812FF4DAC34B2BE CRC64;

Query Match 86.5%; Score 766.5; DB 4; Length 201;  
 Best Local Similarity 77.2%; Pred. No. 4,1e-54;  
 Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2;

QY 4 MNPEYDYLFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDKRTIELDGTIKLQ 63  
 DB 1 MNPEYDYLFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDKRTIELDGTIKLQ 60  
 QY 64 I-----ESFNWVKWMLQEIIDRYASENNKLLVG 91  
 DB 61 IWDTAGGERFRTITSSYRGAGHIIIVYDVTDQESYANVKWMLQEIIDRYASENNKLLVG 120  
 QY 92 NKCDLTKKVVDTYTAKEFADSLGIPLETSAKNATNVEQSFMTAAEIKRMGPGATAG 151  
 DB 121 NKSDLTTRKKVNDNTAKAFADSLGIPLETSAKNATNVEQAFMTAAEIKRMGPGASG 180  
 QY 152 GAESNVRKIOSTPVKQSGGCC 173  
 DB 181 G-ERNLKRIDSTPVKPGAGGCC 201

## RESULT 2

Q9DIG1

ID Q9DIG1 PRELIMINARY; PRT; 201 AA.

AC Q9DIG1 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE 1110011F09RIK protein (RIKEN CDNA 1110011F09 gene).  
 GN 1110011F09RIK  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRIO;  
 RX MEDLINE=21085600; PubMed=11217851;  
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojibori T., Hono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,  
 RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guellinich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmink L.,  
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=KIDNEY;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAS SUBFAMILY.  
 DR EMBL: AK003609; BAB22888.1; -;  
 DR EMBL: BC016408; AAH16408.1; -;  
 DR HSSP: P05713; 3RAB.  
 DR MGD: MGI:1923558; 1110011F09RIK.  
 DR InterPro: IPR003579; GTPase\_Rab.  
 DR InterPro: IPR001806; Ras\_trnsfmg.  
 DR InterPro: IPR002078; Sig54\_interact.  
 DR InterPro: IPR005225; Small\_GTP.

DR Pfam: PF00071; ras; 1.  
 DR PRINTS: PRO0449; RASTRNSFRMG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFAMS: TIGR00231; small\_gtp; 1.  
 DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 KW GTP-binding; lipoprotein.  
 SQ SEQUENCE 201 AA; 22187 MW; 870DF52AEF4B2BE CRC64;

Query Match 85.7%; Score 759.5; DB 11; Length 201;  
 Best Local Similarity 76.2%; Pred. No. 1,5e-53;  
 Matches 154; Conservative 7; Mismatches 8; Indels 33; Gaps 2;

QY 4 MNPEYDYLFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDKRTIELDGTIKLQ 63  
 DB 1 MNPEYDYLFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDKRTIELDGTIKLQ 60  
 QY 64 I-----ESFNWVKWMLQEIIDRYASENNKLLVG 91  
 DB 61 IWDTAGGERFRTITSSYRGAGHIIIVYDVTDQESYANVKWMLQEIIDRYASENNKLLVG 120  
 QY 92 NKCDLTKKVVDTYTAKEFADSLGIPLETSAKNATNVEQSFMTAAEIKRMGPGATAG 151  
 DB 121 NKSDLTTRKKVNDNTAKAFADSLGIPLETSAKNATNVEQAFMTAAEIKRMGPGASG 180  
 QY 152 GAESNVRKIOSTPVKQSGGCC 173  
 DB 181 G-ERNLKRIDSTPVKPGAGGCC 201

## RESULT 3

Q96N61

ID Q96N61 PRELIMINARY; PRT; 141 AA.

AC Q96N61 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE CDNA FLJ33365 fis, clone NB9M4100135, highly similar to RAS-related protein Rab-1A.  
 DE Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Taashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
 RA Fukushima Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,  
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,  
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K.,  
 RA Isoval T.;  
 RT "NEO human cDNA sequencing project.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK055927; BAB71048.1; -;  
 DR InterPro: IPR001806; Ras\_trnsfmg.  
 DR Pfam: PF00071; ras; 1.  
 DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 KW GTP-binding.  
 SQ SEQUENCE 141 AA; 15330 MW; 620F586CA34B3A77 CRC64;

Query Match 79.0%; Score 700; DB 4; Length 141;  
 Best Local Similarity 81.5%; Pred. No. 5,8e-49;  
 Matches 141; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSNPEYDYLFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDKRTIELDGTIKLQ 60  
 DB 1 MSNPEYDYLFKLLIGDSGVGKSCLLRFADPTYESYSTIGVDKRTIELDGTIKLQ 32  
 QY 61 KLOIESFNWVKWMLQEIIDRYASENNKLLVGKCDLTKKVVDTYTAKEFADSLGIPLET 120  
 DB 33 -----ESFNWVKWMLQEIIDRYASENNKLLVGKCDLTKKVVDTYTAKEFADSLGIPLET 88



```

OY 121 TSANKATVWQSPMTMAEIKKRGKPGATGGAENKVIQSTPYKOSGGGCC 173
DB 89 TSANKATVWQSPMTMAEIKKRGKPGATGGAENKVIQSTPYKOSGGGCC 141

RESULT 4
ID 018332 PRELIMINARY; PRT; 205 AA.
AC 018332;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE RAB1 protein.
GN RAB1 OR CG3320.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RX MEDLINE=20196006; PubMed=10731132;
RP STRAIN=BERKELEY.
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sclton G.G., Morten J.R., Tandel M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Arill J.F., Abdayani A., An H.-J., Andrews-Plannkoch C., Balwin D.,
RA Kwan R.H., Doyle C., Baxter E.G., Helt G., Nelson C.W., Miles G.L.G.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu L.B., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster G., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Giodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mileshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murzy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shier E., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Syrakas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Wooley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
RA Ghibs R.A., Myers E.W., Rubin G.M., Ventre J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON R; TISSUE=HEAD;
RX MEDLINE=97228579; PubMed=9074639;
RA Satoh A.K., Tokunaga F., Ozaki K.;
RT "Rab proteins of Drosophila melanogaster: novel members of the Rab-
RT protein family."
RL FEBS Lett. 404:65-69(1997).
CC -1- SIMILARITY: TO RAB PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: AF003734; AAF55873.1;
DR EMBL: D84312; BAA21705.1;

```

---

```

DR HSP: P05713; 3RAB.
DR Flybase: FBgn016700; Rab1.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR002078; S1954_interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras; 1.
DR PRINTS: PR00449; RASTRNSFRNG.
DR SMART: SM00175; RAB; 1.
DR TIGRFRAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 205 AA; 22762 MW; 9CE2691205FAE343 CRC64;

Query Match 77.1%; Score 683; DB 5; Length 205;
Best local similarity 69.4%; Pred. No. 2,2e-47;
Matches 143; Conservative 12; Mismatches 17; Indels 34; Gaps 3;

OY 1 MSSMPEVYTLKLLTGDSGVKSCLLLRPADDTTSTSTIGYDKRTITLDGKTI 60
DB 1 MSSVPEYDYLKLLIGDSGVKSCLLLRPADDTTSTSTIGYDKRTITLDGKTI 60
OY 61 KLOI-----ESNNVKQLQELIDRYASENVKL 88
DB 61 KLOIWPDAQEFRTISYYRGAGIIWYDCITDSEFNKKWLEIEERYACENVNL 120
OY 89 LVGNKCDLTKRKVDYTTAKFADSLGIPLETSAKNATVWQSPMTMAEIKKRGKPGCA 148
DB 121 LVGNKCDLTKRKVDYTTAKFADSLGIPLETSAKNATVWQSPMTMAEIKKRGKPGCA 180
OY 149 TAGAERSNVKI-OSTPYKOSGGGCC 173
DB 181 SAND-NASRVKIDQGRVEMTKSGCC 205

RESULT 5
ID 016143 PRELIMINARY; PRT; 202 AA.
AC 016143;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Small GTP-binding protein.
GN RAB8.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RP SEQUENCE FROM N.A.
RC "Isolation of cDNA encoding small GTP-binding protein from Bombyx
RT mori."
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAB PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: AF013572; AAB67169.1;
DR HSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR002078; S1954_interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras; 1.
DR PRINTS: PR00449; RASTRNSFRNG.
DR SMART: SM00175; RAB; 1.
DR TIGRFRAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 202 AA; 22364 MW; 0338E895063CDPFA CRC64;

Query Match 75.7%; Score 671; DB 5; Length 202;
Best local similarity 66.5%; Pred. No. 2e-46;
Matches 135; Conservative 18; Mismatches 16; Indels 34; Gaps 3;

```

```

OY 4 MNPEVDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRTIELDGTIKLQ 63
DB 1 MNPEVDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRTIELDGTIKLQ 60
OY 64 I-----ESFNNYKWLQEIIDRYASENNKLLVG 91
DB 61 IWDTAGOGERFRTITSSYRGAGHIIYDCTDQDSFSNNKWLQEIIDRYACNNKLLVG 120
OY 92 NKCDDLTKRKVDYTTAKEPADSLGIPLETSAKNATNVQSPMTAAEIKRMGPGATAG 151
DB 121 NKCDDLTKRKVDYTTAKEPADSLGIPLETSAKNATNVQSPMTAAEIKRMGPGATAG 179
OY 152 GAESKNVXI-QSTPVKQSGGGCC 173
DB 180 AAPAGHVKIDQGPIDTGRSCC 202

RESULT 6
OY 09UA06 PRELIMINARY; PRT; 205 AA.
ID 09UA06;
AC 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 22.5 kDa protein.
GN C39F7.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea.
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium."
RL Science 283:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Scheet P., Magg L., Dubbelde C.;
RT "The sequence of C. elegans cosmid C39F7."
RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (Sep-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAS SUBFAMILY.
DB EMBL; AF101310; AAC69218.1; -.
DR HSSP; P01112; 1PLI.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmg.
DR InterPro; IPR002078; S1954_interact.
DR InterPro; IPR005225; Sma11_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; sma11_GTP.1.
DR PROSITE; PS00675; SIGMAS4_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Hypothetical protein; Lipoprotein.
SQ SEQUENCE 205 AA; 22545 MW; 5221920FD05842D3 CRC64;

Query Match 74.5%; Score 660; DB 5; Length 205;
Best Local Similarity 66.2%; Pred. No. 1.5e-45;
Matches 137; Conservative 16; Mismatches 18; Indels 36; Gaps 4;
OY 1 MSNPEVDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRTIELDGTIKL 60
DB 1 MAANPEVDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRTIELDGTIKL 60

```

```

OY 61 KLOI-----ESFNNYKWLQEIIDRYASENNKLLVG 88
DB 61 KLOIWDTAGOGERFRTITSSYRGAGHIIYDCTDQDSFSNNKWLQEIIDRYACNNKLLVG 120
OY 89 LVNKCDDLTKRKVDYTTAKEPADSLGIPLETSAKNATNVQSPMTAAEIKRMGPGCA 148
DB 121 LVNKCDDLTKRKVDYTTAKEPADSLGIPLETSAKNATNVQSPMTAAEIKRMGPGCA 180
OY 149 TAGGAESKNVXI-QSTPVKQSGGGCC 173
DB 181 GAGGA--PEVRIITGSPVQDGRKSGGCC 205

RESULT 7
OY 09HET3 PRELIMINARY; PRT; 201 AA.
ID 09HET3;
AC 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Secretion related GTPase (SrgB).
GN SRGB.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_Taxid=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-N402;
RA Punt P.J., Selboth B., Weenink X.O., van Zeljl C.M., Lenders M.,
Konetchny C., Ram A.F., Montijn R., Kubicek C.P.,
RA Van Den Hondel C.A.;
RT "Identification and characterisation of a family of secretion related
small GTPase encoding genes from the filamentous fungus Aspergillus
niger: a putative SEC4 homologue is not essential for growth."
RL Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAS SUBFAMILY.
DB EMBL; AJ278659; CAC17833.1; -.
DR HSSP; P01112; 1PLI.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmg.
DR InterPro; IPR002078; S1954_interact.
DR InterPro; IPR005225; Sma11_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; sma11_GTP.1.
DR PROSITE; PS00675; SIGMAS4_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 201 AA; 22303 MW; 8DC1A5B0BBD378 CRC64;

Query Match 74.0%; Score 655.5; DB 3; Length 201;
Best Local Similarity 67.2%; Pred. No. 3.4e-45;
Matches 137; Conservative 12; Mismatches 18; Indels 37; Gaps 3;
OY 4 MNPEVDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRTIELDGTIKLQ 63
DB 1 MNPEVDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRTIELDGTIKLQ 60
OY 64 I-----ESFNNYKWLQEIIDRYASENNKLLVG 91
DB 61 IWDTAGOGERFRTITSSYRGAGHIIYDCTDQDSFSNNKWLQEIIDRYACNNKLLVG 120
OY 92 NKCDDLTKRKVDYTTAKEPADSLGIPLETSAKNATNVQSPMTAAEIKRMGPGATAG 151
DB 121 NKCDDLTKRKVDYTTAKEPADSLGIPLETSAKNATNVQSPMTAAEIKRMGPGATAG 179
OY 152 GAESKNVXI-QSTPVKQSGGGCC 173
DB 178 VNNKPTVQVGGQGGQSGSAGGCC 201

```

RESULT 8  
 ID Q9P8J7 PRELIMINARY; PRT: 201 AA.  
 AC Q9P8J7;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE YPLA.  
 GN YPLA.  
 OS Aspergillus awamori.  
 OC Eukaryota; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 NCBI\_TaxID=105351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Saloheimo M., Wang H., Valkonen M., Vasara T., Huuskonen A., Ward M.,  
 RA Penttila M.;  
 RT "The secretory genes ypl1/ypl4 and nsfl/nsf4 from the filamentous  
 RT fungi Trichoderma reesei and Aspergillus niger: evidence for global  
 RT transcriptional regulation of the secretory pathway in Trichoderma  
 RT reesei";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 CC EMBL: AF244543; RefSeq33.1; -.  
 DR HSP: P01112; IPL.  
 DR Interpro: IPR003579; GTPase\_Rab.  
 DR Interpro: IPR001806; Ras\_transfmg.  
 DR Interpro: IPR02078; Sig54\_interact.  
 DR Interpro: IPR05225; Small\_GTP.  
 DR Pfam: PF00071; Ras; 1.  
 DR PRINTS: PRO0449; RASTRNSFRNG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFAMs: TIGR00231; small\_GTP; 1.  
 DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 DR GTP-binding; Lipoprotein.  
 KW GTP-binding; Lipoprotein.  
 SO SEQUENCE 201 AA; 22335 MW; 9D710FBAEBDE279 CRC64;  
 Query Match 73.8%; Score 653.5; DB 3; Length 201;  
 Best Local Similarity 67.2%; Pred. No. 5e-45;  
 Matches 137; Conservative 12; Mismatches 18; Indels 37; Gaps 3;  
 QY 4 MNPEYDYLFKLLIGDSGVGKSCILLRFADPTYESYSTIGVDFKIRITELDGKTIKQ 63  
 DB 1 MNPEYDYLFKLLIGDSGVGKSCILLRFADPTYESYSTIGVDFKIRITELDGKTIKQ 60  
 QY 64 I-----ESPNNVKQMLQEIDRYASENNKILVG 91  
 DB 61 IMDTRAGGERFRITTSYRYGAHGIQVYDVTQDSEFNNVKQMLQEIDRYASENNKILVG 120  
 QY 92 NKCDLTKRYVDYTTAKFADSGIFPLETSKATNTNVEOSFTMAAEIKRKMGFGATAG 151  
 DB 121 NKSDHEKRYVEYTTAKFADSLGIFPLETSKATNTNVEQAFMAAAAEIKRKMA--TAT 177  
 QY 152 GAESNNKIOSPTPVKQSG--GGCC 173  
 DB 178 VNNKPTVQVQGGVQSGSAGCC 201  
 RESULT 9  
 ID Q40203 PRELIMINARY; PRT: 202 AA.  
 AC Q40203;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE RAB1C.  
 GN RAB1C.  
 OS Lotus japonicus.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae; Lotus.  
 NCBI\_TaxID=34305;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE-ROOT NODULES;  
 RX MEDLINE=97231679; PubMed=9076991;  
 RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;  
 RT "Identification of new protein species among 33 different small GTP-  
 RT binding proteins encoded by cDNAs from Lotus japonicus, and expression  
 RT of corresponding mRNAs in developing root nodules";  
 CC Plant J. 11:237-250(1997).  
 RL -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 RC EMBL: 273932; CA98160.1; -.  
 DR HSP: P05713; 3RAB.  
 DR Interpro: IPR003579; GTPase\_Rab.  
 DR Interpro: IPR001806; Ras\_transfmg.  
 DR Interpro: IPR02078; Sig54\_interact.  
 DR Interpro: IPR05225; Small\_GTP.  
 DR Pfam: PF00071; Ras; 1.  
 DR PRINTS: PRO0449; RASTRNSFRNG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFAMs: TIGR00231; small\_GTP; 1.  
 DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 DR GTP-binding; Lipoprotein.  
 KW GTP-binding; Lipoprotein.  
 SO SEQUENCE 202 AA; 22633 MW; 91633EBFAD07662F CRC64;  
 Query Match 72.7%; Score 644; DB 10; Length 202;  
 Best Local Similarity 65.8%; Pred. No. 2.9e-44;  
 Matches 133; Conservative 8; Mismatches 27; Indels 34; Gaps 3;  
 QY 4 MNPEYDYLFKLLIGDSGVGKSCILLRFADPTYESYSTIGVDFKIRITELDGKTIKQ 63  
 DB 1 MNPEYDYLFKLLIGDSGVGKSCILLRFADSDYLDSTYISTIGVDFKIRIVEQDGKTIKQ 60  
 QY 64 I-----ESPNNVKQMLQEIDRYASENNKILVG 91  
 DB 61 IMDTRAGGERFRITTSYRYGAHGIQVYDVTQDSEFNNVKQMLQEIDRYASENNKILVG 120  
 QY 92 NKCDLTKRYVDYTTAKFADSGIFPLETSKATNTNVEOSFTMAAEIKRKMGFGATAG 151  
 DB 121 NKCDLTKRYVDYTTAKFADSGIFPLETSKATNTNVEQAFMAAAAEIKRKMA--SQPN 179  
 QY 152 GAESNNKIOSPTPVKQSG--GGCC 173  
 DB 180 NARPPYQIRGQPVNKR--SGCC 200  
 RESULT 10  
 ID Q9SEH3 PRELIMINARY; PRT: 202 AA.  
 AC Q9SEH3;  
 DT 01-MAR-2000 (TREMBLrel. 13, Created)  
 DT 01-MAR-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE RAS-related small GTP-binding protein (RAS-related small GTP-binding  
 DE protein RAB1C) (At4G17530/DLA800C).  
 GN RAB1C OR AT4G17530.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA; TISSUE=EPIDERMAL;  
 RA Bischoff F., Godde M., Palme J.;  
 RT "Arabidopsis involved in ER-Golgi traffic in meristematic cells in  
 RT Arabidopsis";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Baah J.,

RA Bower L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawada J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Yekker J.R.,  
 RT Arabidopsis cDNA clones."  
 RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAS SUBFAMILY.  
 DR EMBL: AF127134; AAF22133.1; -  
 DR EMBL: AL161546; CAB78756.1; -  
 DR EMBL: AY052204; AKK97675.1; -  
 DR HSSP: P05713; 3RAB.  
 DR InterPro: IPR003579; GTPase\_Rab.  
 DR InterPro: IPR001806; Ras\_trnsfrmng.  
 DR InterPro: IPR002078; Sig54\_interact.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF00071; ras; 1.  
 DR PRINTS: PR00449; RASTRNSFRMNG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFRMS: TIGR00231; small\_GTP; 1.  
 DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 DR GTP-binding; Lipoprotein.  
 KW VARIANT 70 70 F -> L.  
 SQ SEQUENCE 202 AA; 22318 MW; 954FA24C3110FC12 CRC64;  
 Query Match 72.5%; Score 642; DB 10; Length 202;  
 Best Local Similarity 65.8%; Pred. No. 4.2e-44;  
 Matches 133; Conservative 11; Mismatches 24; Indels 34; Gaps 3;  
 QY 4 NMPEYDYLFKLLIDSGVSKCLLRFPADDTYTESYSTIGVDKIRTIELDKTKIKQ 63  
 DB 1 NMPEYDYLFKLLIDSGVSKCLLRFPADDTYTESYSTIGVDKIRTIELDKTKIKQ 60  
 QY 64 I-----ESFNWKQLOEIDRYASFNVKLLVG 91  
 DB 61 IMDTAGQERFRTITSSYRGANGIIVYDVTDLSEFNNKQWLQEI DRYASENVKLLVG 120  
 QY 92 NKCDLTKRKVDYTTAKERFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGATAG 151  
 DB 121 NKCDLTSQKVSTERAKAFADSLGIPLETSAKNATNVEANMATTAIKTMA-SQPRAG 179  
 QY 152 GAESKNVAKI-QSTPY-KOSGGGCC 173  
 DB 180 GSKPPTVQIRGPPVQNO-SGCC 200  
 RESULT 11  
 Q9HDT5 PRELIMINARY; PRT; 202 AA.  
 AC Q9HDT5;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE Small GTP-binding protein YPT1.  
 GN YPT1.  
 OS Trichoderma reesei (Hypocrea jecorina).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 CC Hypocreales; Hypocreaceae; Hypocrea.  
 OX NCBI\_TaxID=51453;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RUIC-30;  
 RA Saloheimo M.L.A., Mann H., Valkonen M., Vasara T., Huuskonen A.,  
 RA Rikonen M., Pakula T., Ward M., Penttilae M.,  
 RT "The secretory genes ypt1/ypt4 and nsf1/nsf4 from the filamentous  
 RT fungi Trichoderma reesei and Aspergillus niger var. awamori: evidence  
 RT for global transcriptional regulation of the secretory pathway in T.  
 RT reesei."  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL: AJ277108; CAC17744.1; -

DR HSSP: P01112; 1PLJ.  
 DR InterPro: IPR003579; GTPase\_Rab.  
 DR InterPro: IPR001806; Ras\_trnsfrmng.  
 DR InterPro: IPR002078; Sig54\_interact.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF00071; ras; 1.  
 DR PRINTS: PR00449; RASTRNSFRMNG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFRMS: TIGR00231; small\_GTP; 1.  
 DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 DR GTP-binding; Lipoprotein.  
 KW SEQUENCE 202 AA; 22421 MW; A60E10E2AB12E7 CRC64;  
 Query Match 72.2%; Score 640; DB 3; Length 202;  
 Best Local Similarity 65.7%; Pred. No. 6.1e-44;  
 Matches 134; Conservative 13; Mismatches 21; Indels 36; Gaps 4;  
 QY 4 NMPEYDYLFKLLIDSGVSKCLLRFPADDTYTESYSTIGVDKIRTIELDKTKIKQ 63  
 DB 1 NMPEYDYLFKLLIDSGVSKCLLRFPADDTYTESYSTIGVDKIRTIELDKTKIKQ 60  
 QY 64 I-----ESFNWKQLOEIDRYASFNVKLLVG 91  
 DB 61 IMDTAGQERFRTITSSYRGANGIIVYDVTDLSEFNNKQWLQEI DRYATEGVKLLVG 120  
 QY 92 NKCDLTKRKVDYTTAKERFADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGATAG 151  
 DB 121 NKSDMSDKRVYTTAKERFADSLGIPLETSAKNATNVEANMATTAIKTMA--GTITA 178  
 QY 152 GAESKNVAKI-QSTPY-KOSGGGCC 173  
 DB 179 NNTKPSVHVQCGQGVNSSSSSC 202  
 RESULT 12  
 Q08154 PRELIMINARY; PRT; 203 AA.  
 AC Q08154;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE GTP-binding protein.  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosida I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RX MEDLINE=94290824; PubMed=8019783;  
 RA Nagano Y., Mural N., Matsuno R., Sasaki Y.,  
 RT "Isolation and characterization of cDNAs that encode eleven small GTP-  
 RT binding protein from Pisum sativum."  
 RL Plant Cell Physiol. 34:447-455(1993).  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL: D12549; BAA02117.1; -  
 DR HSSP: P05713; 3RAB.  
 DR InterPro: IPR003579; GTPase\_Rab.  
 DR InterPro: IPR001806; Ras\_trnsfrmng.  
 DR InterPro: IPR002078; Sig54\_interact.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF00071; ras; 1.  
 DR PRINTS: PR00449; RASTRNSFRMNG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFRMS: TIGR00231; small\_GTP; 1.  
 DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 DR GTP-binding; Lipoprotein.  
 KW SEQUENCE 203 AA; 22463 MW; 6F2E1007A31EF3F4 CRC64;  
 Query Match 71.8%; Score 636.5; DB 10; Length 203;  
 Best Local Similarity 64.9%; Pred. No. 1.2e-43;  
 Matches 131; Conservative 11; Mismatches 27; Indels 33; Gaps 2;

```

OY 4 MNPEDYLFKLLIGSGVSKCLLRFPADDTTSTYISTGVDFKRTIETDGTIKLQ 63
DB 1 MNPEDYLFKLLIGSGVSKCLLRFPADDTTSTYISTGVDFKRTIETDGTIKLQ 60
OY 64 I-----ESFNNVKOMLQEIIDRYASENVNKLTVG 91
DB 61 IMDTAGOERRRTTSSYRGAGHIIIVYDVTDESEFNNVKOMLSEIDRYASDNVKNLLVG 120
OY 92 NKCDLTTRKVVYDTTAKEPADSLGIPLETSAKNATNVEOSFMTMAEIKRKGPGATAG 151
DB 121 NKSDLTARVVSVDITKEFADQIGIPMETSAKDATNVEGAFMAAIAIKRMAQSPSAN 180
OY 152 GAESKNVQIOSFPVKOSGGCC 173
DB 181 NARPPYVQIRGQPVNGOK -GGCC 201

RESULT 13
OY 08M4S8 PRELIMINARY; PRT: 202 AA.
AC 08M4S8:
DB 01-MAR-2002 (TREMblrel. 20, Created)
DB 01-MAR-2002 (TREMblrel. 20, last sequence update)
DB 01-JUN-2002 (TREMblrel. 21, last annotation update)
DB AR4917530/414800.
DB Arabidopsis thaliana (Mouse-ear cross).
DB Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DB Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
DB eurosid II; Brassicales; Brassicaceae; Arabidopsis.
DB NCBI_Taxid=3702;
DB [1]
DB SEQUENCE FROM N.A.
DB Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
DB Bowser L., Canninci P., Chang E., Dale J.M., Goldsmith A.D.,
DB Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlun-Neumann G.,
DB Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
DB Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
DB Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
DB Yamamura Y., Yu G., Shinozaki K., Davis R.W., Theologis A.,
DB Ecker J.R.;
DB "Arabidopsis ORF clones";
DB RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DB EMBL: AY060495; AL31108.1; -.
DB InterPro: IPR003579; GPase_Rab.
DB InterPro: IPR001806; Ras_trnsfmg.
DB InterPro: IPR002078; Slg54_interact.
DB Pfam: PF00071; ras; 1.
DB PRINTS: PR00449; RASTRNSFRMG.
DB SMART: SM00175; Rab; 1.
DB PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
DB GTP-binding, Lipoprotein.
DB SEQUENCE 202 AA; 22286 MW; 954FA12C3110FC12 CRC64;
SO

Query Match 71.6%; Score 634.5; DB 10; Length 202;
Best Local Similarity 65.8%; Pred. No. 1,7e-43;
Matches 131; Conservative 11; Mismatches 24; Indels 33; Gaps 2;

```

```

RESULT 14
OY 09FPJ4 PRELIMINARY; PRT: 202 AA.
AC 09FPJ4:
DB 01-MAR-2001 (TREMblrel. 16, Created)
DB 01-MAR-2001 (TREMblrel. 16, last sequence update)
DB 01-JUN-2002 (TREMblrel. 21, last annotation update)
DB AT5947200 (Putative ras-related small GTP-binding protein).
DB AT5947200/MOL5.5 OR AT5947200.
DB Arabidopsis thaliana (Mouse-ear cross).
DB Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DB Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
DB eurosid II; Brassicales; Brassicaceae; Arabidopsis.
DB NCBI_Taxid=3702;
DB [1]
DB SEQUENCE FROM N.A.
DB Shinozaki K., Chao O., Brooks S., Chen H., Cheuk R., Johnson-Hopson C.,
DB Khan S., Kim C.J., Banh J., Bowser L., Chung M.K., Goldsmith A.D.,
DB Jones T., Karlun-Neumann G., Lam B., Lee J.M., Lin J., Liu S.X.,
DB Miranda M., Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H.,
DB Southwick A., Toriumi M., Yamada K., Yu G., Davis R.W., Theologis A.,
DB Ecker J.R.;
DB "Arabidopsis cDNA clones";
DB RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DB [2]
DB SEQUENCE FROM N.A.
DB Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
DB Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
DB Carlinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
DB Kamiya A., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
DB Seki M., Shinozaki K., Shinozaki K., Davis R.W., Ecker J.R.,
DB Theologis A.;
DB "Arabidopsis full length cDNA clones";
DB RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DB -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAS SUBFAMILY.
DB EMBL: AF080753; AA040342.1; -.
DB HSP: P05713; 3RAB.
DB InterPro: IPR003579; GPase_Rab.
DB InterPro: IPR001806; Ras_trnsfmg.
DB InterPro: IPR002078; Slg54_interact.
DB InterPro: IPR005225; Small_GTP.
DB Pfam: PF00071; ras; 1.
DB PRINTS: PR00449; RASTRNSFRMG.
DB SMART: SM00175; Rab; 1.
DB TRGPRMS: TRG00231; small_GTP; 1.
DB PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
DB GTP-binding, Lipoprotein.
DB SEQUENCE 202 AA; 22313 MW; D0129A296C3EFDPC CRC64;
SO

Query Match 71.4%; Score 633; DB 10; Length 202;
Best Local Similarity 65.8%; Pred. No. 2,2e-43;
Matches 133; Conservative 10; Mismatches 25; Indels 34; Gaps 3;

```

## RESULT 15

024112 ID 024112 PRELIMINARY; PRT; 203 AA.  
 AC 024112;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Small GTP-binding protein.  
 OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4092;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Borisjuk N., Sitallo L., Kaydamov C., Senger S., Tewes A.,  
 RA Manteuffel R.;  
 RT "Cloning and characterization of cDNA clones differentially expressed  
 RT during somatic embryogenesis of Nicotiana plumbaginifolia: a mRNA  
 RT differential display approach."  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL: Y08425; CAA69701.1; -  
 DR HSSP: P01112; IPII.  
 DR InterPro: IPR003579; GTPase\_Rab.  
 DR InterPro: IPR001230; Prenyl\_site.  
 DR InterPro: IPR001806; Ras\_trnsfmg.  
 DR InterPro: IPR002078; Sig54\_interact.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF00071; Ras\_1.  
 DR PRINTS: PR00449; RASTRNSFRMG.  
 DR SMART: SM00175; RAB; 1.  
 DR TIGRFAMs: TIGR00231; small\_gtp; 1.  
 DR PROSITE: PS00294; PRENYLATION; UNKNOWN\_1.  
 DR PROSITE: PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 DR GTP-binding; Lipoprotein.  
 KW SEQUENCE 203 AA; 22516 MW; 13799234PF719DE7 CRC64;  
 SO

Query Match 71.4%; Score 632.5; DB 10; Length 203;  
 Best Local Similarity 63.9%; Pred. No. 2.5e-43;

Matches 129; Conservative 12; Mismatches 28; Indels 33; Gaps 2;

QY 4 MNPEYDYLFLKLLIGDSGVGKSCLLLRADTYTESYSTIGVDFKIRITIELDGKTIKQ 63  
 DB 1 MNPEYDYLFLKLLIGDSGVGKSCLLLRADTYTESYSTIGVDFKIRITIELDGKTIKQ 60  
 QY 64 I-----ESFNNVKOMLOEIDRYASENVNKKLVG 91  
 DB 61 IMDTAGQERPRFTITSSYRGAHGIIVYDVTDOESFNNVKOMLSEIDRYASDSVKKLVG 120  
 QY 92 NKCDLTTKRYVDYTTAFADSLGIFLETSAKNATNVEQSFMTMAAEIKRKGPGATNG 151  
 DB 121 NKCDLTTAKVSTETAFADEIGIPFMTSAKNATNVEQAFMAAASIKRMAASQPSASN 180  
 QY 152 GAESNVKIOSTPYKSGGGCC 173  
 DB 181 NARPPVQIRGQPVNQR-SGCC 201

Search completed: June 18, 2003, 15:23:06  
 Job time : 82 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:21:20 ; Search time 27 Seconds  
(Without alignments)  
188.525 Million cell updates/sec

Title: US-09-820-003b-2

Sequence: 1 MSSNPXYDLFFLLIGDS.....EKSNNKIQSTPWNQSGGSC 173

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database: Issued Patents, AA:\*

- 1: /cgn2\_6/prodata/1/aa/5A-COMB pep:\*
- 2: /cgn2\_6/prodata/1/aa/5B-COMB pep:\*
- 3: /cgn2\_6/prodata/1/aa/6A-COMB pep:\*
- 4: /cgn2\_6/prodata/1/aa/6B-COMB pep:\*
- 5: /cgn2\_6/prodata/1/aa/PCFUS-COMB pep:\*
- 6: /cgn2\_6/prodata/1/aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	813.5	91.8	201	2	US-08-531-525-13 Sequence 13, Appl
2	813.5	91.8	201	2	US-08-718-270A-13 Sequence 13, Appl
3	766.5	86.5	201	4	US-08-916-901-3 Sequence 3, Appl
4	766.5	86.5	201	4	US-08-154-602-3 Sequence 3, Appl
5	760.5	85.8	201	4	US-08-916-901-8 Sequence 8, Appl
6	760.5	85.8	201	4	US-08-154-602-8 Sequence 8, Appl
7	702.5	79.3	202	2	US-08-531-525-14 Sequence 14, Appl
8	702.5	79.3	202	2	US-08-718-270A-14 Sequence 14, Appl
9	439	40.5	215	2	US-08-531-525-10 Sequence 10, Appl
10	439	40.5	215	2	US-08-718-270A-10 Sequence 10, Appl
11	424	47.5	207	2	US-08-531-525-35 Sequence 35, Appl
12	424	47.5	207	2	US-08-718-270A-35 Sequence 35, Appl
13	417.5	47.1	207	2	US-08-531-525-25 Sequence 25, Appl
14	417.5	47.1	207	2	US-08-718-270A-25 Sequence 25, Appl
15	417.5	47.1	207	2	US-08-531-525-51 Sequence 51, Appl
16	417.5	47.1	207	2	US-08-718-270A-51 Sequence 51, Appl
17	390	44.0	198	2	US-08-531-525-34 Sequence 34, Appl
18	390	44.0	198	2	US-08-718-270A-34 Sequence 34, Appl
19	366	41.3	194	2	US-08-531-525-17 Sequence 17, Appl
20	366	41.3	194	2	US-08-718-270A-17 Sequence 17, Appl
21	334	37.7	208	4	US-08-531-525-19 Sequence 19, Appl
22	334	37.7	208	4	US-08-718-270A-19 Sequence 19, Appl
23	330.5	37.3	212	2	US-08-531-525-18 Sequence 18, Appl
24	330.5	37.3	212	2	US-08-718-270A-18 Sequence 18, Appl
25	330.5	37.3	212	2	US-08-531-525-18 Sequence 18, Appl
26	330	37.2	212	2	US-08-718-270A-18 Sequence 18, Appl
27	330	37.2	212	2	US-08-531-525-18 Sequence 18, Appl

28	322	36.3	210	2	US-08-531-525-16	Sequence 16, Appl
29	322	36.3	210	2	US-08-718-270A-16	Sequence 16, Appl
30	319	36.0	214	2	US-08-531-525-52	Sequence 52, Appl
31	319	36.0	214	2	US-08-718-270A-52	Sequence 52, Appl
32	316.5	35.7	190	2	US-08-824-873-3	Sequence 3, Appl
33	316.5	35.7	190	2	US-08-198-184-3	Sequence 3, Appl
34	315.5	35.6	213	2	US-08-531-525-11	Sequence 11, Appl
35	315.5	35.6	213	2	US-08-718-270A-11	Sequence 11, Appl
36	307.5	34.7	203	2	US-08-766-551-8	Sequence 8, Appl
37	298.5	33.7	213	2	US-08-531-525-36	Sequence 36, Appl
38	298.5	33.7	213	2	US-08-718-270A-36	Sequence 36, Appl
39	292.5	33.0	191	4	US-09-075-454-3	Sequence 3, Appl
40	286	32.3	208	2	US-08-531-525-15	Sequence 15, Appl
41	286	32.3	208	2	US-08-718-270A-15	Sequence 15, Appl
42	263.5	29.7	191	2	US-08-531-525-26	Sequence 26, Appl
43	263.5	29.7	191	2	US-08-718-270A-26	Sequence 26, Appl
44	255	28.8	217	2	US-08-773-423-3	Sequence 3, Appl
45	254	28.7	213	2	US-08-773-423-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-08-531-525-13  
Sequence 13, Application US/08531525  
Patent No. 5840683  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5840683, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptides inhibiting the Oncogenic Action  
TITLE OF INVENTION: of p21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Wanner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525  
FILING DATE: 21-Sep-1995  
CLASSIFICATION: 350  
ATTORNEY/AGENT INFORMATION:  
NAME: Feltner, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-531-525-13  
Query Match: 91.8%; Score 813.5; DB 2; Length 201;  
Best Local Similarity 81.2%; Pred. No. 2,3e-87;

Matches 164; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

QY 4 MNPEYDYLFKLLIGDSGVGKSCLLLRPADDTYESYSTIGVDFKIRTIELDGTIKIQ 63  
|||||  
DB 1 MNPEYDYLFKLLIGDSGVGKSCLLLRPADDTYESYSTIGVDFKIRTIELDGTIKIQ 60  
QY 64 I-----ESFNNVKWLOEIDRYASENVKLLVG 91  
|  
DB 61 IWDTAGOERFRITTSYRGAGHILVYDVTDQESFNNVKWLOEIDRYASENVKLLVG 120  
QY 92 NKCDLTTKKVVDYTTAKFPADSLGIPLETSAKNATNVEOSEFMTMAAEIKRMGPGATAG 151  
|||||  
DB 121 NKCDLTTKKVVDYTT-KEFADSLGIPLETSAKNATNVEQAFMTMAAEIKRMGPGATAG 179  
QY 152 GAESKNVKIOSTPVKSGGCC 173  
|:|||||  
DB 180 GSEKSNVNIQSTPVKSGGCC 201

## RESULT 2

US-08-718-270A-13  
; Sequence 13, Application US/08718270A  
; Patent No. 5910478  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.  
; APPLICANT: Plucus, Matthew R.  
; APPLICANT: No. 59104781e, John F.  
; APPLICANT: Abajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptidomimetics Inhibiting  
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P. C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/718.270A  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/531,525  
; FILING DATE: 21-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004,091  
; FILING DATE: 21-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feibet, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 78-95  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 201 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEtical: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Discopyge ommata  
; US-08-718-270A-13

Query Match 91.8%; Score 813.5; DB 2; Length 201;  
Best Local Similarity 81.2%; Pred. No. 2.3e-87;  
Matches 164; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

QY 4 MNPEYDYLFKLLIGDSGVGKSCLLLRPADDTYESYSTIGVDFKIRTIELDGTIKIQ 63  
|||||  
DB 1 MNPEYDYLFKLLIGDSGVGKSCLLLRPADDTYESYSTIGVDFKIRTIELDGTIKIQ 60  
QY 64 I-----ESFNNVKWLOEIDRYASENVKLLVG 91  
|  
DB 61 IWDTAGOERFRITTSYRGAGHILVYDVTDQESFNNVKWLOEIDRYASENVKLLVG 120  
QY 92 NKCDLTTKKVVDYTTAKFPADSLGIPLETSAKNATNVEOSEFMTMAAEIKRMGPGATAG 151  
|||||  
DB 121 NKCDLTTKKVVDYTT-KEFADSLGIPLETSAKNATNVEQAFMTMAAEIKRMGPGATAG 179  
QY 152 GAESKNVKIOSTPVKSGGCC 173  
|:|||||  
DB 180 GSEKSNVNIQSTPVKSGGCC 201

## RESULT 3

US-08-916-901-3  
; Sequence 3, Application US/08916901  
; Patent No. 5892012  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Pretti  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Puvvi  
; TITLE OF INVENTION: RAB PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/916.901  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0367 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-845-4166  
; TELEFAX: 415-855-0555  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 201 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: LIVERUT04  
; CLONE: 2514506  
; US-08-916-901-3

Query Match 86.5%; Score 766.5; DB 2; Length 201;  
Best Local Similarity 77.2%; Pred. No. 7.3e-82;  
Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2;

QY 4 MNPEYDYLFKLLIGDSGVGKSCLLLRPADDTYESYSTIGVDFKIRTIELDGTIKIQ 63



DB 1 MNPEDYLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDGTIKLQ 60  
QY 64 I-----ESPNNVKOMLQIEDRYASENVNKLKLV 91  
DB 61 IMDTAGOERFRTTSSYRGAGIIVYDVTDOESYANVKOMLQIEDRYASENVNKLKLV 120  
QY 92 NKCDLTTKRVYDVTAKFADSLGIPLETSAKNATNVEOSFMTMAAEIKRMGPGATAG 151  
DB 121 NKSDLTTRKRVYDNTAKFADSLGIPLETSAKNATNVEOAFMTMAAEIKRMGPGAAAG 180  
QY 152 GAERKNVKIOSTPVKQSGGCC 173  
DB 181 G-ERPNLKIDSTPVKPGAGGCC 201

RESULT 4  
US-09-154-602-3  
Sequence 3, Application US/09154602  
Patent No. 6300472  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Puryi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,602  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/916,901  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVERMUT04  
CLONE: 2514506  
US-09-154-602-3

Query Match 86.5%; Score 766.5; DB 4; Length 201;  
Best Local Similarity 77.2%; Pred. No. 7.3e-82;  
Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2;  
QY 4 MNPEDYLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDGTIKLQ 63  
DB 1 MNPEDYLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDGTIKLQ 60  
QY 64 I-----ESPNNVKOMLQIEDRYASENVNKLKLV 91  
DB 121 NKSDLTTRKRVYDNTAKFADSLGIPLETSAKNATNVEOAFMTMAAEIKRMGPGAAAG 180

DB 61 IMDTAGOERFRTTSSYRGAGIIVYDVTDOESYANVKOMLQIEDRYASENVNKLKLV 120  
QY 92 NKCDLTTKRVYDVTAKFADSLGIPLETSAKNATNVEOSFMTMAAEIKRMGPGATAG 151  
DB 121 NKSDLTTRKRVYDNTAKFADSLGIPLETSAKNATNVEOAFMTMAAEIKRMGPGAAAG 180  
QY 152 GAERKNVKIOSTPVKQSGGCC 173  
DB 181 G-ERPNLKIDSTPVKPGAGGCC 201

RESULT 5  
US-08-916-901-8  
Sequence 8, Application US/08916901  
Patent No. 5892012  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Puryi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/916,901  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 57006  
US-08-916-901-8

Query Match 85.8%; Score 760.5; DB 2; Length 201;  
Best Local Similarity 76.2%; Pred. No. 3.7e-81;  
Matches 154; Conservative 7; Mismatches 8; Indels 33; Gaps 2;  
QY 4 MNPEDYLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDGTIKLQ 63  
DB 1 MNPEDYLFKLLIGDSGVGKSCLLRFADDTYESYISTIGVDFKRTIELDGTIKLQ 60  
QY 64 I-----ESPNNVKOMLQIEDRYASENVNKLKLV 91  
DB 61 IMDTAGOERFRTTSSYRGAGIIVYDVTDOESYANVKOMLQIEDRYASENVNKLKLV 120  
QY 92 NKCDLTTKRVYDVTAKFADSLGIPLETSAKNATNVEOSFMTMAAEIKRMGPGATAG 151  
DB 121 NKSDLTTRKRVYDNTAKFADSLGIPLETSAKNATNVEOAFMTMAAEIKRMGPGAAAG 180

OY 152 GAESKNVKTOSTPVKSGGCC 173  
| : : : : :  
Db 181 G-ERPNLKIDSTPVKSASGGCC 201

## RESULT 6

US-09-154-602-8  
; Sequence 8, Application US/09154602  
; Patent No. 6300472  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Puri V.  
; TITLE OF INVENTION: RAB PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/154.602  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/916.901  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: P-0367 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 201 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Genbank  
; CLONE: 57006  
; US-09-154-602-8

Query Match 85.8%; Score 760.5; DB 4; Length 201;  
Best Local Similarity 76.2%; Pred. No. 3.7e-81;

Matches 154; Conservative 7; Mismatches 8; Indels 33; Gaps 2;

OY 4 MNPEYDYLFKLLIGDSGVGKSCLLRPADDTYTESYSTIGVDKIRTIELDKGTIKQ 63  
| : : : : :  
Db 1 MNPEYDYLFKLLIGDSGVGKSCLLRPADDTYTESYSTIGVDKIRTIELDKGTIKQ 60  
OY 64 I-----ESFNNVKQWLOEIDRYASENVKLLVG 91  
| : : : : :  
Db 61 IWDTAGOERFRVTSSYYRGAGIIVYDVTDQESYANVKOMLOEIDRYASENVKLLVG 120  
OY 92 NKCDLTFKRVVDYTTAKEFADSLGIPLETSAKNATNVEOSFMTMAAETIKRMGPGATAG 151  
| : : : : :  
Db 121 NKSDLTTRKVVNDNTAKFAFADSLGVPLFETSAKNATNVEOAFMTMAAETIKRMGPGASG 180  
OY 152 GAESKNVKTOSTPVKSGGCC 173  
| : : : : :  
Db 181 G-ERPNLKIDSTPVKSASGGCC 201

## RESULT 7

US-08-531-525-14  
; Sequence 14, Application US/08531525  
; Patent No. 5840683  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.  
; APPLICANT: Pincus, Matthew R.  
; APPLICANT: No. 5840683, John F.  
; APPLICANT: Adajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
; OF P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/531.525  
; FILING DATE: 21-SEP-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 37-94  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 202 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Lymeia stagnalis  
; US-08-531-525-14

Query Match 79.3%; Score 702.5; DB 2; Length 202;  
Best Local Similarity 72.2%; Pred. No. 2.3e-74;

Matches 148; Conservative 9; Mismatches 13; Indels 35; Gaps 4;

OY 1 MSSNPEYDYLFKLLIGDSGVGKSCLLRPADDTYTESYSTIGVDKIRTIELDKGTI 60  
| : : : : :  
Db 1 MSTNPEYDYLFKLLIGDSGVGKSCLLRPADDTYTESYSTIGVDKIRTIELDKGTI 60  
OY 61 KLOI-----ESFNNVKQWLOEIDRYASENVKLL 88  
| : : : : :  
Db 61 KLOIWDTAGOERFRVTSSYYRGAGIIVYDVTDQESFNNVKQWLOEIDRYASENVKLL 120  
OY 89 LVNKCIDLTFKRVVDYTTAKEFADSLGIPLETSAKNATNVEOSFMTMAAETIKRMGPG 148  
| : : : : :  
Db 121 LVNKCIDLTFKRVVDYTTAKEFADSLGIPLETSAKNATNVEOAFMTMAAETIKRMGPG-I 178  
OY 149 TAGGAESKNVKTOSTPVKSGGCC 173  
| : : : : :  
Db 179 TASDS-KPSVKINSSTPSANRGCC 202

## RESULT 8

US-08-718-270A-14  
; Sequence 14, Application US/08718270A

Patent No. 5910478.  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 59104781e, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptidomimetics Inhibiting  
TITLE OF INVENTION: the Oncogenic Action of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/08/718,270A  
APPLICATION NUMBER: 514  
CLASSIFICATION: 514  
PRIORITY APPLICATION DATA:  
PRIORITY APPLICATION NUMBER: US 08/531,525  
FILING DATE: 21-SEP-1995  
PRIORITY APPLICATION DATA:  
PRIORITY APPLICATION NUMBER: US 60/004,091  
FILING DATE: 21-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Felder, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 78-95  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 202 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Lymnea stagnalis  
US-08-718-270A-14

Query Match 79.3%; Score 702.5; DB 2; Length 202;  
Best Local Similarity 72.2%; Pred. No. 2.3e-74;  
Matches 148; Conservative 9; Mismatches 13; Indels 35; Gaps 4;

QY 1 MESANPEVDYLFKLLIGDSGVKSCILLRPADDTYESYFTIGVDPKIRITELDKGTI 60  
DB 1 MGNAMPDDYLFKLLIGDSGVKSCILLRPADDTYESYFTIGVDPKIRITELDKGTI 60  
QY 61 KIQI-----ESFNNTKQWLOEIDRYASENVNKL 88  
DB 61 KQIWDVAGQRRFTTSSYRGAGIIVYDVTDESSENNKQWLOEIDRYASENVNKL 120  
QY 89 LVGNCDLTTKRVVDYTTAKRPAISLGIPELTSARNATNVEOSFTWMAEIKRMGPGA 148  
DB 121 LVGNNSDLYTKR-VDTTAKRPAISLGIPELTSARNATNVEOSFTWMAEIKRMGPG-1 178  
QY 149 TAGGAESKNVYIOSTFVKOSGGCC 173  
DB 179 TASDS-KPSVKNSTPSANKGCC 202

RESULT 9

US-08-531-525-10  
Sequence 10, Application US/08531525  
Patent No. 5840683  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 58406831e, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
TITLE OF INVENTION: of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/08/531,525  
APPLICATION NUMBER: 530  
CLASSIFICATION: 530  
FILING DATE: 21-SEP-1995  
PRIORITY APPLICATION DATA:  
PRIORITY APPLICATION NUMBER: US 08/531,525  
FILING DATE: 21-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Felder, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 215 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Arabidopsis thaliana  
US-08-531-525-10

Query Match 49.5%; Score 439; DB 2; Length 215;  
Best Local Similarity 45.4%; Pred. No. 2e-43;  
Matches 93; Conservative 28; Mismatches 44; Indels 40; Gaps 4;

QY 7 EYDLFLLIGDSGVKSCILLRPADDTYESYFTIGVDPKIRITELDKGTI-- 64  
DB 11 DYDLFLLIGDSGVKSCILLRPADDTYESYFTIGVDPKIRITELDKGTI-- 70  
QY 65 -----ESFNNTKQWLOEIDRYASENVNKL 95  
DB 71 TAGGERTTTAYRGAGIIVYDVTDESSENNKQWLOEIDRYASENVNKL 130  
QY 96 L-TTKRVVDYTTAKRPAISLGIPELTSARNATNVEOSFTWMAEIKRMGPGATAGAE 154  
DB 131 MDESKRAVFPKGOALADEYGIKFFETSAKTNLVEEVEFISGRDIKRLSD--TDSRAE 188  
QY 155 KSNVYIOSTFVKOSG-----GCC 173  
DB 189 PATIKISQTDQAAAGAGATQKSNCC 213

RESULT 10  
US-08-718-270A-10  
Sequence 10, Application US/08718270A  
Patent No. 5910478  
GENERAL INFORMATION:

APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 59104781e, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptidomimetics Inhibiting  
the Oncogenic Action of p21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/718,270A  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/531,525  
FILING DATE: 21-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,091  
FILING DATE: 21-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 78-95  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 215 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Arabidopsis thaliana  
US-08-718-270A-10

Query Match 49.5%; Score 439; DB 2; Length 215;  
Best Local Similarity 45.4%; Pred. No. 2e-43;  
Matches 93; Conservative 28; Mismatches 44; Indels 40; Gaps 4;  
OY 7 EYDYFKLLIGDSGVKSCLLRFADTYTESYSTIGVDPKRTIELDGKTIKQI-- 64  
DB 11 DYDIKLLIGDSGVKSCLLRFSDFTTSTTTGIDPKRTIELDGKRIKQIWD 70  
OY 65 -----ESFNWKQMLQEIIRYASENVNKLIVGNKCD 95  
DB 71 TAGOERFTITAYYNGANGILLVYVTEDESSENNIRNIRNIEQASDNVKNILVGNKAD 130  
OY 96 L-TTKKVVDTYAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRMGPGATAGAAE 154  
DB 131 MDESRAVPTAKGALADEYGIKFPETSAKTLNVEEYFESIGRIDIKQLSD--TDSRAE 188  
OY 155 KSNVYIOSTPVKQSG-----GCC 173  
DB 189 PATIKISQTDAAAGAGATQKSACC 213

RESULT 11  
US-08-531-525-35  
; Sequence 35, Application US/08531525

Patent No. 5840683  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 58406831e, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
of p21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-531-525-35

Query Match 47.9%; Score 424; DB 2; Length 207;  
Best Local Similarity 42.9%; Pred. No. 1.1e-41;  
Matches 82; Conservative 38; Mismatches 41; Indels 30; Gaps 1;  
OY 8 YDYLKLLIGDSGVKSCLLRFADTYTESYSTIGVDPKRTIELDGKTIKQI-- 64  
DB 4 YDLKLLIGDSGVKSCLLRFSDFAFTTSTTTGIDPKRTIELDGKRIKQIWD 63  
OY 65 -----ESFNWKQMLQEIIRYASENVNKLIVGNKCD 97  
DB 64 AGOERFTITAYYNGANGIMKVDITNESFPDKIRNIRNIEHASSDERMILGKCKMN 123  
OY 98 TKKVVDTYAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRMGPGATAGAAEKN 157  
DB 124 EKROYSKERGEKALADYGIKFPETSAKSLNVEEAFITLARDIMTKLNKNSNLSQAV 183  
OY 158 VKIOSTPVKOS 168  
DB 184 DKIKSPKPKPS 194

RESULT 12  
US-08-718-270A-35  
; Sequence 35, Application US/08718270A  
; Patent No. 5910478  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.

```

APPLICANT: No. 59104781e, John F.
APPLICANT: Adajan, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptidomimetics Inhibiting
TITLE OF INVENTION: the Oncogenic Action of p21 Ras
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,270A
FILING DATE: 20-SEP-1996
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/531,525
FILING DATE: 21-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,091
FILING DATE: 21-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fether, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 78-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Discozyme ommata
US-08-718-270A-35

Query Match
Best Local Similarity 47.9%; Score 424; DB 2; Length 207;
Matches 82; Conservative 38; Mismatches 41; Indels 30; Gaps 1;

QY 8 YVYLFKLLIGDSGVKSCLLRPADDTYESTISTIGYDKRTITLDDKTKKQT--- 64
DB 4 YVYLFKLLIGDSGVKSCLLRPADDTYESTISTIGYDKRTITLDDKTKKQT--- 63
QY 65 -----ESNNVQWLOEIDRYASENVNKLKLVGKCDLT 97
DB 64 AQGEFERTIAYYRGAMGIMLVYDITNEKSFDMIRNIRIEHASSDVERMIIGKNCDMN 123
QY 98 TRKVVDTTAKFADSLGIPLETSAKNATNVQSFMTAAEIKRKGPGATAGAEKSN 157
DB 124 ERKQVSKERKGLAIDYGIKFLTSANSSINVEAFITLARDIMTKIKKKNMNSIQEAV 183
QY 158 VTIQSTPYKOS 168
DB 184 DTKKSPKPKS 194

RESULT 13
US-08-824-873-4
; Sequence 4, Application US/08824873
; Patent No. 5843717
; GENERAL INFORMATION:

```

```

APPLICANT: Hillman, Jennifer L.
APPLICANT: Guegler, Karl
TITLE OF INVENTION: NOVEL RAB PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSP for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,873
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PR-0240 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 234746
US-08-824-873-4

Query Match
Best Local Similarity 47.2%; Score 418; DB 2; Length 207;
Matches 84; Conservative 33; Mismatches 47; Indels 32; Gaps 1;

QY 4 MNPXYDLKLLIGDSGVKSCLLRPADDTYESTISTIGYDKRTITLDDKTKKQT 63
DB 1 MNPXYDLKLLIGDSGVKSCLLRPADDTYESTISTIGYDKRTITLDDKTKKQT 60
QY 64 I-----ESNNVQWLOEIDRYASENVNKLKLVG 91
DB 61 IMWYAGERRPRTTAYYRGAMGIMLVYDITNEKSFDMIRNIRIEHASSDVERMIIG 120
QY 92 NKCDLTTRKVVDTTAKFADSLGIPLETSAKNATNVQSFMTAAEIKRKGPGATAG 151
DB 121 NKCDLVNDKQVSKERKGLAIDYGIKFLTSANSSINVEAFITLARDIMTKIKKKN 180
QY 152 GAESKNYKIQSTPYKQ 167
DB 181 SPQSGNQGKTPDQ 196

RESULT 14
US-09-198-184-4
; Sequence 4, Application US/09198184
; Patent No. 6010859
; GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Guegler, Karl
TITLE OF INVENTION: NOVEL RAB PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.

```

STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/198,184  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/824,873  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0240 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 234/46  
US-09-198-184-4

Query Match 47.2%; Score 418; DB 3; Length 207;  
Best Local Similarity 42.9%; Pred. No. 5.5e-41;  
Matches 84; Conservative 33; Mismatches 47; Indels 33; Gaps 1;

QY 4 MNPEYDLEFKLLIDSGVSKSCLLRPADDTYTESYSTIGVDFKRTIELDGKTIKQ 63  
DB 1 MAKTYDLEFKLLIDSGVSKTCVLFRESEDAFNSTFISTIGDFKRTIELDGKRIKQ 60  
QY 64 I-----ESFNNAKOWLOEIDRYASENVNKLTVG 91  
DB 61 IMDTAGOERFTITAYRANGIMLVYDITNEKSPDNIRNMINIEHSAADVERKMIIG 120  
QY 92 NKCDLTTRKVVDTTAKFADSLGIPLETSAKNATNVEOSFTMAAEIKRKGPGATAG 151  
DB 121 NKCDVNDKROVSKERGERKLADYGIKFMETSAKANINVENAFITLARDIKAKDKKLEGN 180  
QY 152 GAESNVKXIQTTPVKQ 167  
DB 181 SPOGSGNGVKITPDQO 196

RESULT 15  
US-08-531-525-25  
Sequence 25, Application US/08531525  
Patent No. 5840683  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5840683le, John F.  
APPLICANT: Adajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
TITLE OF INVENTION: of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder

STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Feiber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 205 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Canis familiaris  
US-08-531-525-25

Query Match 47.1%; Score 417.5; DB 2; Length 205;  
Best Local Similarity 43.5%; Pred. No. 6.2e-41;  
Matches 83; Conservative 33; Mismatches 44; Indels 31; Gaps 1;

QY 8 YVILFKLLIDSGVSKSCLLRPADDTYTESYSTIGVDFKRTIELDGKTIKQ---- 64  
DB 4 YVILFKLLIDSGVSKTCVLFRESEDAFNSTFISTIGDFKRTIELDGKRIKQIMDT 63  
QY 65 -----ESFNNAKOWLOEIDRYASENVNKLTVG 96  
DB 64 AGOERFTITAYRANGIMLVYDITNEKSPDNIRNMINIEHSAADVERKMIIG 123  
QY 97 TTRKVVDTTAKFADSLGIPLETSAKNATNVEOSFTMAAEIKRKGPGATAGAEKS 156  
DB 124 NDKROVSKERGERKLADYGIKFMETSAKANINVENAFITLARDIKAKDKKLEGN 183  
QY 157 NKXIQTTPVKQ 167  
DB 184 NQGVKITPDQO 194

Search completed: June 18, 2003, 15:25:01  
Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 18, 2003, 15:20:05 ; Search time 28 Seconds  
(without alignments)  
668,564 Million cell updates/sec

Title: US-09-820-003b-2

Perfect score: 886

Sequence: 1 MSSMNPEDYLFKLLIGDSGVCSCLLRFADDTYESYISPIGVDRKRTIELDQKTI 173

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubppa/US08\_NEM\_PUB pep:\*  
2: /cgn2\_6/ptodata/2/pubppa/PCT\_NEM\_PUB pep:\*  
3: /cgn2\_6/ptodata/2/pubppa/US06\_NEM\_PUB pep:\*  
4: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/ptodata/2/pubppa/US07\_NEM\_PUB pep:\*  
6: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/2/pubppa/PCTUS\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/2/pubppa/US08\_PUBCOMB pep:\*  
9: /cgn2\_6/ptodata/2/pubppa/US09\_NEM\_PUB pep:\*  
10: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB pep:\*  
11: /cgn2\_6/ptodata/2/pubppa/US10\_NEM\_PUB pep:\*  
12: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/ptodata/2/pubppa/US60\_NEM\_PUB pep:\*  
14: /cgn2\_6/ptodata/2/pubppa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	886	100.0	173	10	US-09-820-003a-2
2	860	97.1	222	10	US-09-820-003a-4
3	768.5	86.7	224	9	US-10-102-806-466
4	766.5	86.5	201	10	US-09-967-736-3
5	760.5	85.8	201	10	US-09-967-736-8
6	444	46.7	207	10	US-09-924-257-8
7	433	46.6	190	10	US-09-822-860-5
8	403.5	45.5	218	10	US-09-925-300-1571
9	390.5	44.1	246	10	US-09-925-302-534
10	371	41.9	162	10	US-09-834-765-766
11	362.5	40.9	198	10	US-09-794-257-16
12	362.5	40.9	198	10	US-09-945-173-5
13	362.5	40.9	198	10	US-09-972-529-4
14	333	36.3	212	10	US-09-350-874-67
15	332	36.3	212	10	US-09-817-198A-4
16	321	36.2	212	10	US-09-817-198A-2
17	321	36.2	212	10	US-09-817-198A-2
18	321	36.2	212	10	US-09-817-198A-2
19	315.5	35.6	223	10	US-09-817-199A-4

20	314.5	35.5	223	10	US-09-817-199A-2	Sequence 2, Appl1
21	313	35.3	146	9	US-09-764-868-698	Sequence 698, App
22	312.5	35.3	222	9	US-09-764-868-1106	Sequence 1106, Ap
23	312.5	35.3	226	9	US-09-764-868-684	Sequence 684, App
24	310	35.0	218	10	US-09-817-198A-5	Sequence 5, Appl1
25	306	34.5	188	9	US-09-764-868-1120	Sequence 1120, Ap
26	292.5	33.0	191	10	US-09-794-257-14	Sequence 14, Appl
27	292.5	33.0	191	12	US-10-051-986-3	Sequence 11, Appl1
28	292.5	33.0	307	9	US-09-764-868-1100	Sequence 1100, Ap
29	292.5	33.0	312	10	US-09-925-302-783	Sequence 783, App
30	292	33.0	59	10	US-09-864-761-35038	Sequence 35038, A
31	267	30.1	216	10	US-09-945-173-10	Sequence 10, Appl
32	267	30.1	222	9	US-09-764-868-1112	Sequence 1112, Ap
33	267	30.1	225	9	US-09-764-868-692	Sequence 692, Appl
34	258	29.1	217	9	US-10-036-542-89	Sequence 89, Appl
35	258	29.1	239	10	US-09-925-301-1077	Sequence 1077, Ap
36	256	28.9	213	9	US-10-036-542-64	Sequence 64, Appl
37	255	28.8	217	10	US-09-988-974-3	Sequence 3, Appl1
38	254	28.7	213	10	US-09-794-257-5	Sequence 5, Appl1
39	254	28.7	213	10	US-09-988-974-8	Sequence 8, Appl1
40	237.5	26.8	208	9	US-10-108-605-45	Sequence 45, Appl
41	235.5	26.6	832	10	US-09-834-765-2	Sequence 2, Appl1
42	235	26.5	217	10	US-09-925-300-1364	Sequence 1364, Ap
43	227	25.6	255	9	US-09-764-868-686	Sequence 686, App
44	219.5	24.8	183	10	US-09-765-298A-26	Sequence 26, Appl
45	219.5	24.8	183	10	US-09-765-298A-28	Sequence 28, Appl

## ALIGNMENTS

RESULT 1  
US-09-820-003a-2  
; Sequence 2, Application US/09820003A  
; Patent No. US20020142382A1  
; GENERAL INFORMATION:  
; APPLICANT: MERRUOV, Genady et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: C1001196  
; CURRENT FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 173  
; TYPE: PRT  
; ORGANISM: Homo sapien  
; US-09-820-003a-2

Query Match 100.0%; Score 886; DB 10; Length 173;  
Best Local Similarity 100.0%; Pred. No. 1.2e-76;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSSMNPEDYLFKLLIGDSGVCSCLLRFADDTYESYISPIGVDRKRTIELDQKTI	60
DB	1	MSSMNPEDYLFKLLIGDSGVCSCLLRFADDTYESYISPIGVDRKRTIELDQKTI	60
QY	61	KLQIESFNVMKQWLOEIDRYASENVKLLVGNKCDLTTRKYVDYTTAREFADSGIPFLE	120
DB	61	KLQIESFNVMKQWLOEIDRYASENVKLLVGNKCDLTTRKYVDYTTAREFADSGIPFLE	120
QY	121	TSKNNNTNVEOSFMTMAAEIKRMKGAGAGAKSNVKIQTSPVKOSGGGCC	173
DB	121	TSKNNNTNVEOSFMTMAAEIKRMKGAGAGAKSNVKIQTSPVKOSGGGCC	173

RESULT 2  
US-09-820-003a-4  
; Sequence 4, Application US/09820003A  
; Patent No. US20020142382A1  
; GENERAL INFORMATION:

APPLICANT: MERKULOV, Gennady et al.  
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
FILE OF INVENTION: PROTEINS, AND USES THEREOF  
CURRENT APPLICATION NUMBER: US/09/820,003A  
CURRENT FILING DATE: 2001-03-29  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 222  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-820-003A-4

Query Match 97.1%; Score 860; DB 10; Length 222;

Best Local Similarity 84.4%; Pred. No. 5e-74; Mismatches 0; Indels 32; Gaps 1;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNPEYDYLKLLIGDSGVKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 60  
DB 18 MSSNPEYDYLKLLIGDSGVKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 77  
QY 61 KIQI-----ESFNNVKQWLOEIDRYASENVNKL 88  
DB 78 KIQIDTAGQERFRTITSSYRGAGIIVYDVTDOESFNNVKQWLOEIDRYASENVNKL 137  
QY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKRMGPAA 148  
DB 138 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKRMGPAA 197  
QY 149 TAGGAEKSNVKIQSTPVKOSGGGCC 173  
DB 198 TAGGAEKSNVKIQSTPVKOSGGGCC 222

## RESULT 3

US-10-102-806-466  
Sequence 466, Application US/10102806  
Publication No. US2003005442A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA103PICI  
CURRENT APPLICATION NUMBER: US/10/102,806  
CURRENT FILING DATE: 2002-03-22  
PRIOR APPLICATION NUMBER: 09/925,298  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05881  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 846  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 466  
LENGTH: 224  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-102-806-466

Query Match 86.7%; Score 768.5; DB 9; Length 224;

Best Local Similarity 76.5%; Pred. No. 2.5e-65; Mismatches 7; Indels 33; Gaps 2;

Matches 156; Conservative 8; Mismatches 7; Indels 33; Gaps 2;

QY 2 SSANPEYDYLKLLIGDSGVKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 61  
DB 22 SSANPEYDYLKLLIGDSGVKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 81  
QY 62 LQI-----ESFNNVKQWLOEIDRYASENVNKL 89  
DB 82 LQIMTAGQERFRTITSSYRGAGIIVYDVTDOESYANVKQWLOEIDRYASENVNKL 141  
QY 90 VGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKRMGPAT 149

DB 142 VGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKRMGPAA 201  
QY 150 AGGAESNVKIQSTPVKOSGGGCC 173  
DB 202 SGG-ERPMLKIDSTPVKPPAGGCC 224

## RESULT 4

US-09-967-736-3  
Sequence 3, Application US/09967736  
Patent No. US20020103340A1  
GENERAL INFORMATION:  
APPLICANT: Hallman, Jennifer L.  
Ial, Preeti  
Corley, Neil C.  
Shah, Purvi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/967,736  
FILING DATE: 28-Sep-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/154,602  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: L1YRTUT04  
CLONE: 2514506

SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-967-736-3

Query Match 86.5%; Score 766.5; DB 10; Length 201;

Best Local Similarity 77.2%; Pred. No. 3.3e-65; Mismatches 7; Indels 33; Gaps 2;

Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2;

QY 4 MNPEYDYLKLLIGDSGVKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 63  
DB 1 MNPEYDYLKLLIGDSGVKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 60  
QY 64 I-----ESFNNVKQWLOEIDRYASENVNKL 91  
DB 61 IWDTAGQERFRTITSSYRGAGIIVYDVTDOESYANVKQWLOEIDRYASENVNKL 120  
QY 92 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKRMGPATAG 151  
DB 121 NKSDLTTKKVVDYTTAKFADSLGIPLETSAKNATNVQSFMTMAAEIKRMGPAGAS 180  
QY 152 GAESNVKIQSTPVKOSGGGCC 173



Mon Jun 23 16:54:10 2003

us-09-820-003b-2.rapb

Page 3

DB 181 G-ERNKLTSTPYKPSAGGCC 201

RESULT 5  
US-09-967-736-8  
Sequence 8, Application US/09967736

Patent No. US20020103340A1  
GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.  
Corley, Neil C.  
Shah, Puryl

TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA

COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION NUMBER: US/09/967,736  
FILING DATE: 28-Sep-2001  
PRIORITY APPLICATION DATA:  
PRIORITY APPLICATION NUMBER: 09/154,602

ATTORNEY/AGENT INFORMATION:  
FILING DATE: <unknown>  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 57006  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-967-736-8

Query Match 85.8%; Score 760.5; DB 10; Length 201;  
Best Local Similarity 76.2%; Pred. No. 1,2e+64;

Matches 154; Conservative 7; Mismatches 8; Indels 33; Gaps 2;

DB 61 IMDTAGGERPRTYSSYRGAGHIIYVDYDQESYAVKWLQEDIRYASENVKLLVG 91

DB 92 NNCOLTRKYVDYTTKEFADSLGIPLETSAKNATNVEOSFTMAEIKRMGATAG 151

DB 121 NMSDLTKKYVDMTAKFADSLGVPLETSAKNATNVEOAFMTMAEIKRMGPGASG 180

DB 152 GAENSNYKIOSTPYKOSGGCC 173

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

DB 181 G-ERNKLTSTPYKPSAGGCC 201

RESULT 6  
US-09-794-257-8  
Sequence 8, Application US/09794257

Patent No. US2002009804A1  
GENERAL INFORMATION:

APPLICANT: Meyers, Rachel  
TITLE OF INVENTION: 32705, 23224, 32700, 32712, No. US2002009804A1

FILE REFERENCE: 32705, 23224, 32700, 32712, No. US2002009804A1

CURRENT APPLICATION NUMBER: US/09/794,257  
FILING DATE: 2001-02-27

PRIORITY APPLICATION DATA:  
PRIORITY APPLICATION NUMBER: 60/185,606  
FILING DATE: 2000-02-29

SOFTWARE: FASTSEQ for Windows Version 4.0  
NUMBER OF SEQ ID NOS: 16

SEQ ID NO 8  
LENGTH: 207

ORGANISM: homo sapiens

US-09-794-257-8

Query Match 46.7%; Score 414; DB 10; Length 207;  
Best Local Similarity 44.2%; Pred. No. 1e-31;

Matches 84; Conservative 31; Mismatches 37; Indels 38; Gaps 2;

DB 61 IMDTAGGERPRTYSSYRGAGHIIYVDYDQESYAVKWLQEDIRYASENVKLLVG 91

DB 92 NNCOLTRKYVDYTTKEFADSLGIPLETSAKNATNVEOSFTMAEIKRMGATAG 151

DB 121 NMSDLTKKYVDMTAKFADSLGVPLETSAKNATNVEOAFMTMAEIKRMGPGASG 180

DB 146 PGATAGCAEK 155

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

DB 181 NSAGAGGPVK 190

OY 97 TTKKVDYTTAKFADSLGIPLETSAKNATNVEQSFTMAAEIKRMGPGATAGAEKS 156  
 Db 121 NEKROVSERGERKLAIDYGIKIFLETSAKSSINVEAFITLARDIMTKIKKNENSIGEA 180  
 OY 157 NVKIOSTPVK 166  
 Db 181 VDKLSPPK 190

## RESULT 8

US-09-925-300-1571  
 ; Sequence 1571, Application US/09925300  
 ; Patent No. US20020151681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Ruben  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101  
 ; CURRENT APPLICATION NUMBER: US/09/925,300  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1890  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 1571  
 ; LENGTH: 218  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-925-300-1571

Query Match 45.5%; Score 403.5; DB 10; Length 218;  
 Best Local Similarity 39.7%; Pred. No. 1,1e-30;  
 Matches 85; Conservative 32; Mismatches 30; Indels 67; Gaps 4;

OY 8 YDYLKLLIGDSGVKSCLLIRPADDTYTESYSTIGVDKIRKTIETDGTIKIQT--- 64  
 Db 24 YDLFKLLIGDSGVKCTVLRFSDDAFNTFTSTIGIDKIKIVELQGRKIKIQTWT 83  
 OY 65 -----ESFNVKOMLOEIDRYASBNVKNKLGVNKCND 95  
 Db 84 AGQERHHTTTSYRGAMGIMLYDITNGKSPENISKLRNIDELANDVERMLIGNKCND 143  
 OY 96 LTTKRVVDTTAKFADSLGIPLETSAKNATNVEQSFTMAAEIKRMGPGATAGAEK 155  
 Db 144 MDDRKYVKGKGEOLAREHGIRFETSAKANNINIEKAFLLAEDI----- 188  
 OY 156 SNVKIOSTPVKQ-----SGGS-----CC 173  
 Db 189 ---LRKTPVKEPSENVDISSGGGVGWRKSKCC 218

## RESULT 9

US-09-925-302-534  
 ; Sequence 534, Application US/09925302  
 ; Patent No. US2002004941A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA104  
 ; CURRENT APPLICATION NUMBER: US/09/925,302  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05918  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 896  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 534  
 ; LENGTH: 246  
 ; TYPE: PRT

; ORGANISM: Homo sapiens  
 US-09-925-302-534

Query Match 44.1%; Score 390.5; DB 10; Length 246;  
 Best Local Similarity 41.7%; Pred. No. 2.2e-29;  
 Matches 78; Conservative 35; Mismatches 41; Indels 33; Gaps 2;

OY 1 MSSNPEYDYLKLLIGDSGVKSCLLIRPADDTYTESYSTIGVDKIRKTIETDGTIKI 60  
 Db 41 VSAMAKAYDHLFKLLIGDSGVKCTVLRFSDDAFNTFTSTIGIDKIKIVELQGRKIKI 100  
 OY 61 KLOI-----ESFNVKOMLOEIDRYASBNVKNKL 88  
 Db 101 KLOVWDYTAGOERFHTTAYRGAMGILLYDITNGKSPENISKLRNIDELANDVERMLIGNKCND 160  
 OY 89 LVGNKCDLTTKRVVDTTAKFADSLGIPLETSAKNATNVEQSFTMAAEIKRMGPGATAGAEK 148  
 Db 161 ILGNKCDKAKRKVKQEKQADKLAREHGIRFETSAKSSNANVDEAFSSLIARDILKSG-GR 219  
 OY 149 TAGAEK 155  
 Db 220 RSGNGNK 226

## RESULT 10

US-09-834-765-766  
 ; Sequence 766, Application US/09834765  
 ; Patent No. US20020055478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Pia M. Chailita-Eld  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Daniel E.H. Afar  
 ; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
 ; FILE REFERENCE: AND DETECTION OF CANCER  
 ; CURRENT FILING DATE: 2001-09-21  
 ; PRIOR APPLICATION NUMBER: 60/197,647  
 ; PRIOR FILING DATE: 2000-04-12  
 ; NUMBER OF SEQ ID NOS: 770  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 766  
 ; LENGTH: 162  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-834-765-766

Query Match 41.9%; Score 371; DB 10; Length 162;  
 Best Local Similarity 45.1%; Pred. No. 9e-28;  
 Matches 73; Conservative 28; Mismatches 29; Indels 32; Gaps 1;

OY 12 FKLLIGDSGVKSCLLIRPADDTYTESYSTIGVDKIRKTIETDGTIKIQT----- 64  
 Db 1 FKLLIGDSGVKCTVLRFSDDAFNTFTSTIGIDKIKIVELQGRKIKIQTWTAGOE 60  
 OY 65 -----ESFNVKOMLOEIDRYASBNVKNKLGVNKCNDLTK 99  
 Db 61 RFRITTAAYRGAMGIMLYDITNGKSPENISKLRNIDELANDVERMLIGNKCNDVNDK 120  
 OY 100 KVDYTTAKFADSLGIPLETSAKNATNVEQSFTMAAEIKRMGPGATAGAEK 141  
 Db 121 ROVSERGERKLAIDYGIKIFLETSAKANNINIEKAFLLAEDI----- 162

## RESULT 11

US-09-794-257-16  
 ; Sequence 16, Application US/09794257  
 ; Patent No. US20020009804A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Meyers, Rachel

```

; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009604A1el
; FILE REFERENCE: Human G-Proteins
; CURRENT APPLICATION NUMBER: US/09/794,257
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam accession number PF00071
US-09-794-257-16

```

```

Query Match 40.9%; Score 362.5; DB 10; Length 198;
Best Local Similarity 41.7%; Pred. No. 7,4e-27;
Matches 86; Conservative 25; Mismatches 40; Indels 55; Gaps 6;

```

```

QY 13 KLLIGDSGVKSCLLIRPADDTYESYISTGVDFKRTIELDGKTIKLOI----- 64
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 KVLIGDSGVKSCSLIRPTDNKFVEEYIPTIGVDYTKTVEDGKTVKLIQIMDTAGGER 60
QY 65 -----ESFNVKOMLOEIDRYA--SENVNKLIVGNKCDLT- 97
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 FRALRPAYYRGAGFLLVYDITSRDSFENYKMKLEILRHADKDENVPIVYGNKCDLED 120
QY 98 -----TKRYVDYTTAKEFADSLG-IPLETSAKNAVNEOSFMTMAEIKRRMGPG 147
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 DEDLELEGKRVYSTEGBEALAKELGALPMEISAKNTINVEAFELAREILIKVY--- 177
QY 148 ATAGAERSNVKIQSTPVKOSGGGCC 173
DB 178 -----SEVNVNLDQ-PARKKRSKCC 196

```

```

RESULT 12
US-09-945-173-5
; Sequence 5, Application US/09945173
; Patent No. US20020127568A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 47324, A NOVEL HUMAN G-PROTEIN AND USES
; FILE REFERENCE: 38155-20035.00
; CURRENT APPLICATION NUMBER: US/09/945,173
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,293
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-945-173-5

```

```

Query Match 40.9%; Score 362.5; DB 10; Length 198;
Best Local Similarity 41.7%; Pred. No. 7,4e-27;
Matches 86; Conservative 25; Mismatches 40; Indels 55; Gaps 6;
QY 13 KLLIGDSGVKSCLLIRPADDTYESYISTGVDFKRTIELDGKTIKLOI----- 64
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 KVLIGDSGVKSCSLIRPTDNKFVEEYIPTIGVDYTKTVEDGKTVKLIQIMDTAGGER 60
QY 65 -----ESFNVKOMLOEIDRYA--SENVNKLIVGNKCDLT- 97
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 FRALRPAYYRGAGFLLVYDITSRDSFENYKMKLEILRHADKDENVPIVYGNKCDLED 120

```

```

QY 98 -----TKRYVDYTTAKEFADSLG-IPLETSAKNAVNEOSFMTMAEIKRRMGPG 147
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 DEDLELEGKRVYSTEGBEALAKELGALPMEISAKNTINVEAFELAREILIKVY--- 177
QY 148 ATAGAERSNVKIQSTPVKOSGGGCC 173
DB 178 -----SEVNVNLDQ-PARKKRSKCC 196

```

```

RESULT 13
US-09-972-529-4
; Sequence 4, Application US/09972529
; Patent No. US20020150916A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 47316, A NOVEL HUMAN G-PROTEIN AND USES
; FILE REFERENCE: 38155-20041.00
; CURRENT APPLICATION NUMBER: US/09/972,529
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/237,716
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-972-529-4

```

```

Query Match 40.9%; Score 362.5; DB 10; Length 198;
Best Local Similarity 41.7%; Pred. No. 7,4e-27;
Matches 86; Conservative 25; Mismatches 40; Indels 55; Gaps 6;
QY 13 KLLIGDSGVKSCLLIRPADDTYESYISTGVDFKRTIELDGKTIKLOI----- 64
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 KVLIGDSGVKSCSLIRPTDNKFVEEYIPTIGVDYTKTVEDGKTVKLIQIMDTAGGER 60
QY 65 -----TKRYVDYTTAKEFADSLG-IPLETSAKNAVNEOSFMTMAEIKRRMGPG 147
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 FRALRPAYYRGAGFLLVYDITSRDSFENYKMKLEILRHADKDENVPIVYGNKCDLED 120
QY 98 -----TKRYVDYTTAKEFADSLG-IPLETSAKNAVNEOSFMTMAEIKRRMGPG 147
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 DEDLELEGKRVYSTEGBEALAKELGALPMEISAKNTINVEAFELAREILIKVY--- 177
QY 148 ATAGAERSNVKIQSTPVKOSGGGCC 173
DB 178 -----SEVNVNLDQ-PARKKRSKCC 196

```

```

RESULT 14
US-09-350-874-67
; Sequence 67, Application US/09350874
; Patent No. US20020019020A1
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; TITLE OF INVENTION: METHODS FOR TREATING CARDIOVASCULAR DISORDERS
; FILE REFERENCE: NMT-069
; CURRENT APPLICATION NUMBER: US/09/350,874
; PRIOR FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 22, 2003, 18:32:10 ; Search time 138 seconds  
(without alignments)  
1839.601 Million cell updates/sec

Title: US-09-820-003b-2

Sequence: 1 MSNMPEYDYFLKLLIGDS.....EKSNTKIQSPYKSGGSGCC 173

Scoring table:

BIOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Pgapop 6.0, Pgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL-frame.p2n.model -DEV-xip  
-O/cg92.1/USPRO.spool/US09820003/unat.18062003.144132.13246/app.query.fasta.1.327  
-DB-published\_applications\_NA -OFMT-fastap -SUFFIX-p2n.rnpb -MINMATCH-0.1  
-LOOPCL-0 -LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-biosum62  
-TRANS-human40.cdi -LIST-45 -DOCALIGN-200 -THR SCORE-pct -THR MAX-100  
-THR MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pco -NORM-ext -HEA5T5IE-500 -MINLEN-0  
-MAXLEN-200000000 -USER-US09820003.ecgn.1.1.80.ernat.18062003.144132.13246  
-NCPU-6 -ICPU-3 -NO MMAP -LARGEROVERY -NRG SCORES-0 -WAIT -DSBPLOCK-100  
-LONGLOG -DEV TIMEOUT-120 -WARN TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5  
-FGAPOP-6 -FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database: Published Applications\_NA:

- 1: /cg92.6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cg92.6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cg92.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cg92.6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cg92.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cg92.6/ptodata/2/pubpna/PCT05\_PUBCOMB.seq.\*
- 7: /cg92.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cg92.6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cg92.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cg92.6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cg92.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cg92.6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cg92.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cg92.6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	886.100	1405	10	US-09-820-003A-1
2	768.5	86.7	935	US-09-967-736-4
3	768.5	86.7	935	US-10-102-806-48
4	642	72.5	609	US-09-938-842A-832

5	619.5	69.9	881	10	US-09-770-445-529	Sequence 529, App
6	487.5	55.0	46050	10	US-09-820-003A-3	Sequence 3, Appl
7	481	54.3	601	10	US-09-820-003A-36	Sequence 36, Appl
8	428.5	48.4	651	9	US-09-938-842A-836	Sequence 836, Appl
9	414	46.7	624	10	US-09-794-257-9	Sequence 9, Appl
10	414	46.7	1161	10	US-09-794-257-9	Sequence 7, Appl
11	414	46.7	2497	10	US-09-834-975-879	Sequence 879, Appl
12	414	46.7	2497	10	US-09-834-975-885	Sequence 885, App
13	414	46.7	2497	10	US-09-834-975-885	Sequence 885, App
14	414	46.7	2497	10	US-09-834-975-885	Sequence 885, App
15	403.5	45.5	1537	10	US-09-925-302-91	Sequence 91, Appl
16	390.5	44.1	1274	10	US-09-925-302-91	Sequence 91, Appl
17	364	41.1	549	9	US-09-925-302-91	Sequence 91, Appl
18	344	38.6	538	10	US-09-925-302-91	Sequence 91, Appl
19	343	38.7	898	10	US-09-925-302-91	Sequence 91, Appl
20	337.5	38.1	1069	9	US-09-925-302-91	Sequence 91, Appl
21	333	37.6	639	10	US-09-925-302-91	Sequence 91, Appl
22	332	37.5	487	10	US-09-925-302-91	Sequence 91, Appl
23	329.5	37.2	2021	9	US-09-925-302-91	Sequence 91, Appl
24	329.5	37.1	459	9	US-09-925-302-91	Sequence 91, Appl
25	327.5	37.0	1479	10	US-09-822-830A-365	Sequence 365, App
26	327.5	36.9	585	10	US-09-817-198A-1	Sequence 1, Appl
27	326.5	36.9	585	10	US-09-917-800A-1461	Sequence 1461, App
28	324	36.6	427	10	US-09-960-352-8911	Sequence 8911, App
29	324	36.6	483	9	US-09-918-995-2704	Sequence 2704, App
30	323	36.5	833	9	US-09-918-995-2704	Sequence 2704, App
31	322	36.3	457	9	US-09-918-995-2704	Sequence 2704, App
32	318	35.9	374	10	US-09-878-574-3637	Sequence 3637, App
33	318	35.9	654	10	US-09-878-574-3637	Sequence 3637, App
34	316.5	35.7	894	10	US-09-770-445-487	Sequence 487, App
35	316	35.7	3936	10	US-09-770-445-487	Sequence 487, App
36	314.5	35.5	875	12	US-10-051-986-10	Sequence 10, Appl
37	314.5	35.5	1106	9	US-10-051-986-10	Sequence 10, Appl
38	314.5	35.5	1116	10	US-10-037-270-959	Sequence 959, Appl
39	314.5	35.5	2674	10	US-09-794-257-13	Sequence 13, Appl
40	313	35.3	442	9	US-09-817-199A-1	Sequence 1, Appl
41	312.5	35.3	1316	9	US-09-764-868-85	Sequence 85, Appl
42	312.5	35.3	1316	9	US-09-764-868-85	Sequence 85, Appl
43	309	34.9	2623	9	US-09-764-868-85	Sequence 85, Appl
44	309	34.9	446	9	US-09-764-868-85	Sequence 85, Appl
45	306	34.5	566	9	US-09-764-868-85	Sequence 85, Appl

## ALIGNMENTS

RESULT 1  
US-09-820-003A-1  
; Sequence 1, Application US/09820003A  
; Patent No. US20020142382A1  
; GENERAL INFORMATION:  
; APPLICANT: MERKIOV, Gennady et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; FILE REFERENCE: C1001196  
; CURRENT APPLICATION NUMBER: US/09/820,003A  
; CURRENT FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1405  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-820-003A-1

Alignment Scores:  
Pred. No.: 2.33e-104  
Score: 886.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 10  
Length: 1405  
Matches: 173  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-820-003b-2 (1-173) x US-09-820-003a-1 (1-1405)

```

QY 1 MetSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
DB 180 ATGTCAGCAGCATGAAATCCGGAATATGATATTTATTCAGATTACTTCTGATTTGGCAGACA 239
QY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspPheThrTyrThrGluSerTyr 40
DB 240 GGGGTGGAAAGCTGCTGCTCTCTTGTGAGTTGAGATGATACATATACAGAAAGCTAC 299
QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
DB 300 ATCAGCAGCAATGTTGTTGATGATTTCAAAATAGAACTATAGCTTATGACGGGAATACATC 359
QY 61 LysLeuGluIleGluSerPheAsnAsnValLysGluIlePheGluGluIleAspArgTyr 80
DB 360 AAGCTTCAATATAGAGTCTCTCAATATATGTAAACAGTGGTGCAGGAATATGATGCTTAT 419
QY 81 AlaSerGluAsnValAsnLysLeuLeuValGlyAsnLysCysAspPheThrThrLysLys 100
DB 420 GCCAGTGAATATGTCACAAATTTGTTGAGGAACAAATGTGATCTACACAAAGAAA 479
QY 101 ValValAspTyrThrThrAlaLysGluPheAlaAspSerLeuGlyIleProPheLeuGlu 120
DB 480 GTAGTAGCTACACACAGCAGCAAGAAATTTGCTGATTCCTTGGAAATTCGGTTTGGAA 539
QY 121 ThrSerAlaLysAsnAlaThrAsnValGluGluSerPheMetThrMetAlaAlaGluIle 140
DB 540 ACCAGTGTAAAGATGCAACGAATGTAGACAGCTTTTCATGACAGATGCGAGCTGAGATT 599
QY 141 LysLysArgMetGlyProGlyAlaThrAlaGlyLysGluLysSerAsnValLysIle 160
DB 600 AAAAGCCAAAGGGTCCCGGAGCAACAGCTGCTGCTGAGAAAGTCCAAATGTTAAAAAT 659
QY 161 GlnSerThrProValLysGlnSerGlyLysCys 173
DB 660 CAGAGCACTCCAGTCAAGCACTCAGGTGAGGTGCTCTC 698

```

# RESULT 2

US-09-967-736-4  
 : Sequence 4, Application US/09967736  
 : Patent No. US2002010340A1

## GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.  
 Lal, Preeti  
 Corley, Neil C.  
 Shah, Purvi  
 TITLE OF INVENTION: RAB PROTEINS  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Dr.  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: PASTESQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/967,736  
 FILING DATE: 28-Sep-2001  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/154,602  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0367 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 925 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: LIVERUT04

CLONE: 2514506

US-09-967-736-4  
 : SEQUENCE DESCRIPTION: SEQ ID NO: 4:

## Alignment Scores:

Pred. No.:	1.78e-89	Length:	925
Score:	768.50	Matches:	156
Percent Similarity:	80.39%	Conservative:	8
Best Local Similarity:	76.47%	Mismatches:	7
Query Match:	86.74%	Indels:	33
DB:	10	Gaps:	2

US-09-820-003b-2 (1-173) x US-09-967-736-4 (1-925)

```

QY 2 SerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGly 21
DB 60 GCCGCGCATGAACCCCGAATATGACTTACTGTTTAAAGCTGCTTTGATTTGGCGACTCAGGC 119
QY 22 ValGlyLysSerCysLeuLeuLeuArgPheAlaAspPheThrTyrThrGluSerTyrIle 41
DB 120 GTGGGCAAGTATGCTGCTCTCTCTGCGGTTTCTGATGACAGTACAGAGAGCTTACTC 179
QY 42 SerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys 61
DB 180 AGCAGCATCGGGGTGAGTCTCAAGATCCGAGACCATCGAGCTGATGGCAAAATATATCAA 239
QY 62 LengthIle----- 64
DB 240 CTTCAGATCTGGACACAGCGGGCCAGAGACGGTTCGGACCATCACTTCAGCTACTAC 299
QY 65 -----GluSerPheAsnAsn 69
DB 300 CGGGGGGCTATGCGATCATCGTGTGATGACGCTACGACAGGAATCTTACGCCAAC 359
QY 70 ValLysGluIlePheGluGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 89
DB 360 GTGAAGCAGTGGCTGACAGAGATTGACCGCTATGCCAGCGAAGCTCAATTAAGCTCTG 419
QY 90 ValGlyAsnLysCysAspPheThrThrLysLysValValAspTyrThrThrAlaLysGlu 109
DB 420 GTGGGCAACAGAGCGACTTACACACAGAAAGGTGGAGAACACACACAGCCAGAGAG 479
QY 110 PheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnVal 129
DB 480 TTTCAGACTCTCGGGGATCCCTCTTGTGAGAGAGCGCCAGGAATGCCACCAATGTC 539
QY 130 GluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThr 149
DB 540 GAGCAGCGCTTCATGACCATCGCTGCTGTAATAAAGGAGGGGCTCGAGAGAGCC 599
QY 150 AlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 169
DB 600 TCTGGGGGCG--GAGCGGGCCCAATCTCAAGATGACAGACACCCCTGTAAAGCCGGTGC 656
QY 170 GlyGlyCysCys 173
DB 657 GGTGCTGTTC 668

```

## RESULT 3

US-10-102-806-48  
 : Sequence 48, Application US/10102806  
 : Publication No. US20030054421A1

## GENERAL INFORMATION:

APPLICANT: Rosen et al.



```
DB 481 GGTTCATGGCCATGCTCTGCATCAACCAAGACGATGCG---AGCCACCTGCAGCA 537
OY 152 GYALAGLysSerValLysIleGlnSerThrProValLysGlnSerGlyGly 171
DB 538 GGAATCAAGCCACCAAGGCTCCAGATCCAGACAACTGTATACCAAGCA---TCAGGC 594
OY 172 CysCys 173
DB 595 TCCTGC 600

RESULT 5
US-09-770-445-529
; Sequence 529, Application US/09770445
; Patent No. US2002023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krieger, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurdan, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 202305 (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 529
; LENGTH: 881
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-529

Alignment Scores:
Pred. No.: 2,71e-70 Length: 881
Score: 619.50 Matches: 125
Percent Similarity: 68.97% Conservative: 15
Best Local Similarity: 61.58% Mismatches: 30
Query Match: 69.92% Indels: 33
DB: 10 Gaps: 2

US-09-820-003b-2 (1-173) x US-09-770-445-529 (1-881)
OY 3 SerMetAsnProGlnIleuArgPheAlaAspThrTyrThcGlnSerTyrIleSer 22
DB 57 ACCATGAAATCCGAGTACACTATCTTTTCAAGCTCTGCTATCGGGGATTCGCGCTA 116
OY 23 GlyLysSerCysLeuLeuLeuArgPheAlaAspThrTyrThcGlnSerTyrIleSer 42
DB 117 GCGAAGCTCTGCTCTTCTTTGAGATTCCTGATTCATTAAGTAAAGTTACATTAC 176
OY 43 ThrIleGlyValAspPheLysIleArgThrIleGlnLeuAspGlyLysThrIleLysLeu 62
DB 177 ACTAATGAGATGCGATTTTAAATATAGGACTGTGAACAAGATGCGAAACATTAAGC 236
OY 63 GlnIle----- 64
```

```
DB 237 CAAATTTGGGACACTGCTGTGTCAGAACGCTTCAGAGCTATTACTAGACTTACCCT 296
OY 65 -----GlnSerPheAsnVal 70
DB 297 GGGGACATGAAATATTATTGTCTACAGATGTCACAGATGAAGAACTTCATATATGTC 356
OY 71 LysGlnTrpLeuGlnGlnIleAspArgTyrAlaSerGluAsnValAsnLysLeuVal 90
DB 357 AAGCAATGGTTGAGTAAATGTATGCTGTATGCTAGTACAAATGTCACAACTCCCTGTT 416
OY 91 GlyAsnLysCysAspLeuThrThrLysValValAspTyrThrAlaLysGlnPhe 110
DB 417 GGAACACAGCTGATCTTCTGATGAAACACAGGACATCTTATGAACCTCCCAAGCTTTT 476
OY 111 AlaAspSerLeuGlyIleProPheLeuGlnTrpSerAlaLysAsnAlaThrAsnValGlu 130
DB 477 GCGATGAATAGGGATCTCTTTATGAGACTAGTGAAGATGCTACAAACGTAGCA 536
OY 131 GlnSerPheMetThrMetAlaAlaGlnIleLysLysArgMetGlyProGlyAlaThrAla 150
DB 537 CAGGCTTCATGGCAATGTCGTGCATCCATCAAGAGAGATGCTAGCCACACGCTGGC 596
OY 151 GlyGlyAlaGlnLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 170
DB 597 AATAATGCAAGACCAACCGACGCTGCAGATCAGAGACAGCTGTGACACAGAG---AAC 653
OY 171 GlyCysCys 173
DB 654 GCGTCTGC 662

RESULT 6
US-09-820-003a-3
; Sequence 3, Application US/09820003A
; Patent No. US20020142382A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001196
; CURRENT APPLICATION NUMBER: US/09/820,003A
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 46050
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(46050)
; OTHER INFORMATION: n = A,T,C or G
US-09-820-003a-3

Alignment Scores:
Pred. No.: 1,03e-50 Length: 46050
Score: 487.50 Matches: 113
Percent Similarity: 51.07% Conservative: 6
Best Local Similarity: 48.50% Mismatches: 16
Query Match: 55.02% Indels: 99
DB: 10 Gaps: 2

US-09-820-003b-2 (1-173) x US-09-820-003a-3 (1-46050)
OY 24 LysSerCysLeuLeuLeuArgPheAlaAspThrTyrThcGlnSerTyrIleSerThr 43
DB 42782 AGCTAGCAAGACGATGTTCTGTATTTCACAGACTATATATATATATATATACAGC 42841
OY 44 IleGlyValAspPheLysIleArgThrIleGlnLeuAspGlyLysThrIleLysLeuGln 63
DB 42842 TTTCAAATGCTCTTT----- 42856
```



```
OY 64 IIEGUSerPheasnsanValysGlnTrpLeuGlnLulleasparGlyralaserGlu 83
DB 42857 TAGGAGTCCTCAATATGTAAACAGTGGCTCAGAAATAGATCCGTATGCCAGTGA 42916
OY 84 AsnValAsnLysLeuValGlyAsnLysCysAspLeuThrThrLysValValAsp 103
DB 42917 AATGTCAACAAATGTTGGTAGGAAACAAATGATCTACACCAAGAAAGTAGTAGNC 42976
OY 104 TTTTThrAlaLys----- 108
DB 42977 TACAAACAGGAA--GGTATGTTTAAAGTTTATTTTCATCTGATTTGAAGGTGTA 43035
OY 108 ----- 108
DB 43036 ATTATGTATGGGTTTCGACGTAAACGTAAAGCCACTTTTAAATATGTCACATAG 43095
OY 108 ----- 108
DB 43096 AATTACTGTGACAGTGCATTTGTGTAGCATCTGTTGGATCCAAATGACTTAGTCCCTC 43155
OY 108 ----- 108
DB 43156 ACGCTCATATGAGATGTAGAAATGCAATGAAATTAAGTAAAGATTTTTCAGTGT 43215
OY 109 ----- 109
DB 43216 AATTGTGCTCATATTTCTTCTTATGAAATTTGCTGATTCCTTGGAAATTCGGTTTGGAA 43275
OY 121 ThrSerAlaLysAsnAlaThrAsnValGluGlnSerPheMetThrMetAlaLagLulle 140
DB 43276 ACCAGTGTAAAGATCAACAGCATGTAGAACATCTTTCATGACAGATGCGAGTGGAT 43335
OY 141 LysLysArgMetGlyProGlyAlaThrAlaGlyLysLysSerAsnValLysLys 160
DB 43336 AAAAAGCAATGGGTCCCGAGCAACAGCTGGTGGTGGAGAGTGCATGTTAAAT 43395
OY 161 GlnSerThrProValLysGlnSerGlyGlyCysCys 173
DB 43396 CAGACACTCCAGTCAAGCAGTGCAGTGGTGTGC 43434

RESULT 7
US-09-820-003A-36
: Sequence 36, Application US/09820003A
: Patent No. US20020142382A1
: GENERAL INFORMATION:
: APPLICANT: MERKIOV, Gennady et al.
: TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
: FILE REFERENCE: C1001196
: CURRENT APPLICATION NUMBER: US/09/820,003A
: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 36
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Homo sapien
US-09-820-003A-36

Alignment Scores:
Pred. No.: 1,12e-52 Length: 601
Score: 481.00 Matches: 107
Percent Similarity: 56.02% Conservative: 0
Best Local Similarity: 56.02% Mismatches: 1
Query Match: 54,29% Indels: 84
DB: 10 Gaps: 1

US-09-820-003B-2 (1-173) x US-09-820-003A-36 (1-601)
OY 66 SerPheAsnValLysGlnTrpLeuGlnLulleasparGlyralaserGluAsnVal 85
DB 3 TCCCTCAATATGTTTAAACAGTGGCTCAGAAATAGATCTGTTGACAGTGAATATGTC 62
```

```
OY 86 AsnLysLeuValGlyAsnLysCysAspLeuThrThrLysValValAspLysThr 105
DB 63 AACAAATTTGGTGGGACAAATGATGTGACCAACAAAGAAATAGTACAGTACACA 122
OY 106 ThrAlaLys----- 108
DB 123 ACACGAA--GGTATGTTTAAAGTTTATTTTCATCTGATTTGAAGGTGTTGAATATG 161
OY 108 ----- 108
DB 162 TATGGGTTTCGACGTAAACGTAAAGCCACTTTTAAATATGTCATGAAATACT 241
OY 108 ----- 108
DB 242 GTGACGTACAAATTTGTGTAGCATCTGTTGGATCCAAATGACTAGTCCACGCTY 301
OY 108 ----- 108
DB 302 CATATGATGTAGAAATGCAATGAAATTAAGTAAAGATTTTTCAGTGTAAATGT 361
OY 109 ----- 109
DB 362 GCTCATATATCTCTTAAAGATTTGCTGATTCCTTGGAAATTCGGTTTGGAAACAGT 421
OY 123 AlaLysAsnAlaThrAsnValGluGlnSerPheMetThrMetAlaLagLulleLysLys 142
DB 422 GCTAAGATGCAACCAATGTAGAACATCTTTCATGACATGCGACCTGAGTTAAAAAG 481
OY 143 ArgMetGlyProGlyAlaThrAlaGlyLysLysSerAsnValLysLys 162
DB 482 CGAATGGTCCCGAGCAACAGCTGGTGGTGTGAGAAATTCATTTAAATTCACAGC 541
OY 163 ThrProValLysGlnSerGlyGlyCysCys 173
DB 542 ACTCAGTCAAGCATCAGTGAAGGTGTCTCTC 574

RESULT 8
US-09-938-842A-836
: Sequence 836, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPI300-3
: CURRENT APPLICATION NUMBER: US/09/938,842A
: NUMBER OF SEQ ID NOS: 24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: LENGTH: 651
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-836

Alignment Scores:
Pred. No.: 7,41e-46 Length: 651
Score: 428.50 Matches: 92
Percent Similarity: 58.83% Conservative: 31
Best Local Similarity: 44.02% Mismatches: 39
Query Match: 48,36% Indels: 47
DB: 9 Gaps: 5

US-09-820-003B-2 (1-173) x US-09-938-842A-836 (1-651)
```

```

QY 7 GluYrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyValGlyLysSerCys 26
DB 31 GATACGATTAACCTATTAACCTCTGATCGAGCAGACGCGCTGTGTAAGATTC 90
QY 27 LeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyrIleSerThrIleGlyVal 46
DB 91 CTTCCTTACGTTTCTTCAGATGCTGCTTACCACCGATTTTCAACATTTGGGAT 150
QY 47 AspPheLysIleArgThrIleGlyLeuAspGlyLysThrIleLysLeuGlnIle 64
DB 151 GATTTTAAGATACGAGCTATTGAGCTCGATGGAGAGATTAAGCTGCAATCTCGGAT 210
QY 64 ----- 64
DB 211 ACTGCCGACAGAGCGGTTCCGCAATCACAACCTGCTACTACCGTGGAGCANTGGG 270
QY 65 -----GluSerPheAsnAsnValLysGlnTyrPleu 74
DB 271 ATTTGCTGTGTATGANGTACATGATGATCATCTTTCACACATCGAGGATTTGGATC 330
QY 75 GlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuValGlyLysLysCys 94
DB 331 CGTACATTTGACGACAGCTCTCTGATGCTCACACAGATTTCTAGTTGGACAAAGCA 390
QY 95 Aspleu---ThrThrLysLysValAlaAspTyrThrThrAlaLysGluPheAlaAspSer 113
DB 391 GATATGATGANAAGAAAGACCTGTGCCAATAATCTAAGGCCAAGCTCTGCAGATGAA 450
QY 114 LeuGlyIleProPheLysGluThrSerAlaLysAsnAlaThrAsnValGluGlnSerPhe 133
DB 451 TATGAAATGAACTTTTTCGAGCTAGTGTGCCAAGACTTAACCTTAAGGAAAGTTTTC 510
QY 134 MetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThrAlaGlyAla 153
DB 511 TTCTGTATTCCTTAAGACATTAAGCAAGACTTGCAGAT-----ACCATGACGACGCT 564
QY 154 GluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGlyGly----- 171
DB 565 GAGCGCGCAAAACATCAAAATCAACCAATCC-----GACCAAGGTGCGGGAAACATCT 615
QY 172 -----CysCys 173
DB 616 CAAGCTACTCAGAAATCAGCATGTTGC 642

```

RESULT 9  
 US-09-794-257-9  
 ; Sequence 9, Application US/09794257  
 ; Patent No. US20020009804A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Meyers, Rachel  
 ; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1  
 ; FILE REFERENCE: 35800/209285  
 ; CURRENT APPLICATION NUMBER: US/09/794,257  
 ; PRIOR FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: 60/185,606  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 624  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 ; US-09-794-257-9

Alignment Scores:  
 Pred. No.: 5,16e-44 Length: 624  
 Score: 414.00 Matches: 87  
 Percent Similarity: 60.50% Conservative: 34  
 Best Local Similarity: 43.50% Mismatches: 41  
 Query Match: 46.73% Indels: 39  
 DB: 10 Gaps: 3

```

US-09-820-003b-2 (1-173) x US-09-794-257-9 (1-624)
QY 4 MetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyValGly 23
DB 1 ATGGCGACAGCGTATGATTAATCTTCAAGCTCTGCTGATGCGGAGCTGGGGTAGGC 60
QY 24 LysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyrIleSerThr 43
DB 61 AAGACCTGCTCCTGTTCCGCTTTCAGAGGAGCGGCTTAACACACCTTATCTCCACC 120
QY 44 IleGlyValAspPheLysIleArgThrIleGlyLeuAspGlyLysThrIleLysLeuGln 63
DB 121 ATCGGAATGATTTAAATTAAGAACGATAGAACGAAAGAAATTAAGCTTCAG 180
QY 64 Ile----- 64
DB 181 ATATGGGACACAGCGGGGTGAGAAAGATTCCGAACATCAGACAGCTACTACAGAGA 240
QY 65 -----GluSerPheAsnAsnValLys 71
DB 241 GCGATGGCATTAATGCTGTATGATCATCACAATGAAATTCCTTGACATATTAATA 300
QY 72 GlnTyrPheGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuValGly 91
DB 301 AATTGATTCAGAAACATTCAGACATGCTCTCCGATGCGAAGAAATGATCCTGGGT 360
QY 92 AsnLysCysAspLeuThrThrLysLysValAlaAspTyrThrThrAlaLysGluPheAla 111
DB 361 AACCAATGTGATATGATGACAAACAGCAAGTGTCAAAAGAGGGGAGGAGCTAGCA 420
QY 112 AspSerIleGlyIleProPheLysGluThrSerAlaLysAsnAlaThrAsnValGluGln 131
DB 421 ATTGACTATGGAATTAATTTCTTGAGACAAACGCAAAATCCAGTCAATGTACAGAG 480
QY 132 SerPheMetThrMetAlaAlaGluIle-----LysLysArgMetGlyProGly 147
DB 481 GCATTTTTCACCTTCGACGAGATTAATGACAAACCTCAACAGAAATGATGACAGC 540
QY 148 AlaThrAlaGlyValGluLysSerAsnValLysIleGlnSerThrProValLysGln 167
DB 541 AATTGACGAGGACGAGT-----GGACCAGTGAATAATACAGAAA-CCGATCAAGAA 593

```

RESULT 10  
 US-09-794-257-7  
 ; Sequence 7, Application US/09794257  
 ; Patent No. US20020009804A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Meyers, Rachel  
 ; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1  
 ; FILE REFERENCE: 35800/209285  
 ; CURRENT APPLICATION NUMBER: US/09/794,257  
 ; PRIOR FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: 60/185,606  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 1161  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (18)...(641)  
 ; US-09-794-257-7

Alignment Scores:  
 Pred. No.: 1.3e-43 Length: 1161  
 Score: 414.00 Matches: 87  
 Percent Similarity: 60.50% Conservative: 34  
 Best Local Similarity: 43.50% Mismatches: 41  
 Query Match: 46.73% Indels: 39

```

DB: 10 Gaps: 3
US-09-820-003B-2 (1-173) x US-09-794-257-7 (1-1161)

OY 4 MetasnrProclumLyAspPrlrLeuPhelLysLeuLeuIleGlyAspSerGlyValGly 23
Db 18 ATGGCGAAGACGACGACATTCCTTCTTCAAGCTCTGCTGATGGCGAGCTGGGGGTAGC 77
OY 24 LysSerCysLeuLeuLeuArgPheAlaAspAspThrTyThrGlnSerTyrlIleSerThr 43
Db 78 AAGACGCGCTCTCTGCTCCGCTTCGAGAGAGCGCTTCAACACACCGCTTCATCTCCACC 137
OY 44 IleGlyValAspPheLysIleArgThrIleGlnLeuAspGlyLysThrIleLysLeuGln 63
Db 138 ATCGGAATTCATTTAAATAATTGAAACGATAGAACTAGATGGAAGAATAATTAACTTCAG 197
OY 64 Ile----- 64
Db 198 ATATGGGACACAGCGGCTCAGAAAGATTCGACACATCAGACAGAGCTACTACAGACA 257
OY 65 -----GlnSerPheAsnValLys 71
Db 258 GCCATGGCGCATATGCGTGCTGTATGACATACACAAATGAAATAATCCTTGACATTTTAA 317
OY 72 GlnPrlLeuGlnGluLeaAspArgTyAlaSerGluAsnValAsnLysLeuValGly 91
Db 318 AATTGGATTCGAAACATTGGAAGACGATGCTCTCCGATGTCGAAAGAATGATCTCGGT 377
OY 92 AsnLysCysAspLeuThrThrLysLysValValAspTyThrThrAlaLysGluPheAla 111
Db 378 AACAAATGTAATATGATGACAAAAGCAAGGTCTCAAAAAGAGAGGAGGAGAACTACGA 437
OY 112 AspSerLeuGlyIleProPheLeuGlnThrSerAlaLysAsnAlaThrAsnValGluGln 131
Db 438 ATTGACTATGCGATTAAATTCCTGAGACAAAGCGCAAAATCCAGTCCAAATGTAAGAG 497
OY 132 SerPheMetThrMetAlaIleGluIle-----LysLysArgMetGlyProGly 147
Db 498 GCATTTTTCACCTTGACGACGAGATATATGACAAATCTCAACAGAAATGAATGACAGC 557
OY 148 AlaThrIleGlyIleGlyLysSerAsnValLysIleGlnSerThrProAlaLysGln 167
Db 558 AATTCACACAGAGACAGCGT-----GAGCACAGTGAAATATACAGAAA-CCGATCAAGAA 610

RESULT 11
US-09-834-975-879
: Sequence 879, Application US/09834975
: Patient No. US20020110815H1
GENERAL INFORMATION:
APPLICANT: Lilile, James
APPLICANT: Brown, Jeffrey
APPLICANT: Bolt, Andrew
APPLICANT: Van Hufel, Christophe
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
OF HUMAN CANCERS
FILE REFERENCE: MRI-016B
CURRENT APPLICATION NUMBER: US/09/834,975
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/197,538
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 1046
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 879
LENGTH: 2497
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(2497)
OTHER INFORMATION: n = A..T..C or G
US-09-834-975-879

```

Alignment Scores:	4.06e-43	Length:	2497
Pred. No.:	414.00	Matches:	87
Score:	60.508	Conservative:	34
Percent Similarity:	43.508	Mismatches:	41
Best Local Similarity:	46.738	Indels:	39
Query Match:	10	Gaps:	3

US-09-820-003B-2 (1-173) x US-09-834-975-879 (1-2497)

OY	4	MeCAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyValGly	23
	11		
Db	92	ATGGCGAAGACGACGATTAATCTTCAAGCTCCGCGATCGGCACACTCGGGGGTGAAGC	131
	24	LysSerCysLeuLeuLeuAArgPheAlaAspThrTyrThrGluSerTyrIleSerThr	43
	152	AAGACCGCGCTCCCTGCTCCGCTTCACAGAGAGCGCCCTTCAACACCCACTTCATCTCCACC	211
OY	44	IleGlyValAspPheLysIleArgThrIleGlyLeuAspGlyTyrThrIleLysLeuGln	63
Db	212	ATCGCAATGATTTTAAAAATTAGAACGATAGAACTAGATGAGAAAGAAATTTAGCTTCAG	271
OY	64	Ile-----	64
Db	272	ATATGGGACACAGCGGCTCAGAAAGATTCCGACATCCAGACGACGACTAGACAGGA	331
OY	65	-----GluSerPheAsnValLys	71
Db	332	GCCATGGCGCATTAATGCTGCTATGATCACTCAAAATGAAAATCTTTACACTATTAA	391
OY	72	GlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuValGly	91
	1392		
Db	392	AATTGGATCGAAACATTGAAGACATGCCCTCTCCATGTCACAAAGATGACTCGGGT	451
OY	92	AsnLysAspSerLeuThrThrLysValValAspTyrThrAlaLysGluPheAla	111
	453		
Db	453	AACAAATGTAATGATGACACAAAGCAAGTGTCAAAAGAAAGGGAGAGACTAGCA	511
OY	112	AspSerLeuGlyTyrProPheLysGluThrSerAlaLysAsnAlaThrAsnValGluGln	131
Db	512	ATTGACTATGGGATTAATTTCTTGGAGACAAAGCGCAAAATTCAGTCAAAATGTGAAGAG	571
OY	132	SerPheMetAlaIleGluIle-----LysLysArgMetGlyProGly	147
	572	GCATTTTTCACCTTGACGACGAGATATAAGACAAAACTCAACAGAAAATGTAAGACAGC	631
OY	148	AlaThrAlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGln	167
Db	632	AATTCACGACGAGGACGAGT-----GGACCAAGTGAATAATACAGAAAA-CCGATCAAGAA	684

RESULT 12  
 US-09-834-975-885  
 : Sequence 885, Application US/09834975  
 : Patent No. US20020110815A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Lillie, James  
 : APPLICANT: Brown, Jeffrey  
 : APPLICANT: Bolt, Andrew  
 : APPLICANT: Van Huffel, Christophe  
 : TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS  
 : FILE REFERENCE: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
 : FILE REFERENCE: WTI-0158  
 : CURRENT FILING DATE: US/09/834,975  
 : PRIOR APPLICATION NUMBER: 2001-04-13  
 : NUMBER OF SEQ ID NOS: 1046  
 : SOFTWARE: fastseq for Windows Version 4.0  
 : SEQ ID NO 888  
 : LENGTH: 2497  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-885

```

Alignment Scores:	
Pred. No.:	4, 06e-43
Score:	41.00
Percent Similarity:	60.50%
Best local Similarity:	43.50%
Query Match:	46.73%
DB:	10
Length:	249
Matches:	87
Conservative:	34
Mismatches:	41
Indels:	39
Gaps:	3

US-09-820-003B-2 (1-173) x US-09-834-975-885 (1-2497)

```

OY      4 MetanProGluIuYrAsPTrYLleuPhelysLeuLeuLlEglYasPerGIyValIGly 23
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      92 ATGGCGAAGACGTATGCATTTCTCTTCACAGCTCCTGCTGATGGCCACTCGGGGATGCC 151
OY      24 LysSerCysLeuLeuLeuArGpheAlaAspSPThrTYrrHcGlusErTyrILleSerThr 43
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      152 AAGCACTGCGCTCCGTTCCTCTTAGAGAGCGCTTCACACACCCTTCATCTCCACC 211
OY      44 ILleGIyValaSpPheLYsIIeaRgTHrlleGluLeuAspGLYseHrILleLyseuGln 63
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      212 ATCGGAATTGATTTTAATAATTAGACATGAACTGAACTGAGAAGAAAATTAACTTCAG 271
OY      64 Ile----- 64
Db      272 ATATGGGACACAGCGGGTCAAGAAAGATTCCGAACAATCACGACGCTACTACAGAGA 331
OY      65 ----- GlusPheasnValys 71
        :|||||:|||||:
Db      332 GCCATGGCGCATTTGCTGCTGCTATGCATCATCAAATGAAGAAAATCCCTTGCAATATTAA 391
OY      72 GlnTrPheugInGluIleaSPArGYrAlaSeRGIuaSnValaSnLYsLeuLeuValGIy 91
        |||:|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||
Db      392 AATTGGATCAGAAACATTTGAAGACATGCCCTCTTCGAGTGCAGAAAGATGATCCCGGT 451
OY      92 AsnLYsCysAspleuThrThrLYsLYsValValaSPTrYrrHrAlalysGluPIeAla 111
        |||:|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||
Db      452 AACCAAAAGTATATGAAATGACAAAAGCAAGTGTCAAAAAGAAAGAGGGAGGAAGTCAGCA 511
OY      112 AspSerLeuLYsLIlePrOpheLeuGluThrSerAlLYsnaIlathraSnValIGluGln 131
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      512 ATTGCATCATGGGATTAATTCCTTGACACACAGCCCAAAATCCAGTCCAATGTAGAAAGC 571
OY      132 SerPheNethrMetaLaalaGluIle----- LysLYsArgMetGLYProGIy 147
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      572 GCATTTTTTACACTYGACGAGATATATGACAAAACCTCAACAGAAAATAATGATGCACG 631
OY      148 AlaThrLaclGLyGlaaGluLYsSerASnValLYsILleGlnSerThrProValLYsGln 167
        :|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||
Db      632 AATTCACGACGAGCAGGT-----GGACGAGTGAATAATACAGAAAA-CCGATCAAAAGAA 684

RESULT 13
US-09-834-975-894
Sequence 894, Application US/09834975
Patent No. US20020110815A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Brown, Jeffrey
APPLICANT: Bolt, Andrew
APPLICANT: Van Hufel, Christophe
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
FILE REFERENCE: MRI-016B
CURRENT APPLICATION NUMBER: US/09/834,975
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/197,538
PRIOR FILING DATE: 2000-04-14
```

```

: NUMBER OF SEQ ID NOS: 1046
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 894
:
: LENGTH: 2497
:
: TYPE: DNA

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-894

```

Alignment Scores:	
Pred. No.:	4,06e-43
Score:	414.00
Percent Similarity:	60.50
Best local Similarity:	43.508
Query Match:	46.738
DB:	10
Length:	2497
Matches:	87
Conservative:	34
Mismatches:	41
Indels:	39
Gaps:	3

US-09-820-003B-2 (1-173) x US-09-834-975-894 (1-2497)

[illegible]

```

1  TITLE OF INVENTION:  OF HUMAN CANCERS
2
3  FILE REFERENCE:  MRI-016B
4
5  CURRENT APPLICATION NUMBER:  US/09/834,975/58
6
7  CURRENT FILING DATE:  2001-04-13
8
9  PRIOR APPLICATION NUMBER:  60/197,538
10
11  PRIOR FILING DATE:  2000-04-14
12
13  NUMBER OF SEQ ID NOS:  1046
14
15  SOFTWARE:  FastSeq for Windows Version 4.0.0
16
17  SEQ ID NO 896
18
19  LENGTH:  2497
20
21  TYPE:  DNA
22
23  ORGANISM:  Homo sapiens
24
25  FEATURE:
26
27  NAME/KEY:  misc_feature
28
29  LOCATION:  (1)..(2497)
30
31  OTHER INFORMATION:  n - A,T,C or G
32
33  US-09-834-975-896

```

Alignment Scores:	
pred. No.:	4,06e-43
Score:	44.00
Percent Similarity:	60.50%
Best Local Similarity:	43.50%
Query Match:	46.75%
DB:	10
	Gaps:
	3
	Length:
	2497
	Matches:
	87
	Conservative:
	34
	Mismatches:
	41
	Indels:
	39
	Gaps:
	3

[illegible]

```

1  APPLICANT:  Craig Rosen,
2  APPLICANT:  Steve Ruben
3  TITLE OF INVENTION:  Nucleic Acids, Proteins and Antibodies
4  FILE REFERENCE:  P1101
5  CURRENT APPLICATION NUMBER:  US/09/925,300
6  CURRENT FILING DATE:  2001-08-10
7  PRIOR APPLICATION NUMBER:  PCT/US00/05988
8  PRIOR FILING DATE:  2000-03-08
9  PRIOR APPLICATION NUMBER:  60/124,270
10 PRIOR FILING DATE:  1999-03-12
11 NUMBER OF SEQ ID NOS:  1890
12 SOFTWARE:  PatentIn Ver. 2.0
13 SEQ ID NO 631
14 LENGTH:  1537
15 TYPE:  DNA
16 ORGANISM:  Homo sapiens
17 FEATURE:
18 NAME/KEY:  misc feature
19 LOCATION:  (5)
20 OTHER INFORMATION:  n equals a,t,g, or c
21 US-09-925-300-631

```

Alignment Scores:	4,46e-42	Length:	1537
Prod. No.:		Matches:	85
Score:	403.50	Conservative:	32
Percent Similarity:	54.67%	Mismatches:	30
Best Local Similarity:	39.72%	Indels:	67
Query Match:	45.54%	Gaps:	4
DB:	10		

US-09-820-003B-2 (1-173) x US-09-925-300-631 (1-1537)

OY	8	TYASPTYRLPEHLYSLEUENLEILGSLYASPSERGLYALGLYASERCYLEU	27
Db	548	TACGACCTGCTTTTCACAGCTGCTCTGATCGGGGATTCGGAGTGGGGAAACCTGCGC	607
OY	28	LEULEAATGPPHEALASPSAPPTHTYRTHGLUSERTYRILLESERTHILGLYALAP	47
Db	608	CTTTTCGTTTTGGGTATGCTCTCAATACACTTTATTTCCACCATGAGTAATGAC	667
OY	48	PHELYILEATGTHRIELIULEUASPSLYSTHRIILEYSEUHLLE-----	64
Db	668	TTCAAGATCAAAACAGTGTGAATTAATCAAGAGAAAGAAATCAAGCTACAGATGGGTATCA	727
OY	64	-----	64
Db	728	GCAGGCCAGAGGAGATTTCACACATCAACCTCTACTACAGAGGCCAATGGGTATC	787
OY	65	-----	75
Db	788	ATGCTAGTATATGACATCAACCAATGTGAATGTTGAAACATCGCAATGGCTTAA	847
OY	76	GLUILLAPAAATGYRIALASERGLUASPSALASUYSLEUVAIGLYASUICYASAP	95
Db	848	AAATATGATGACCTGCCATATMAAGTGTGAAGAAATGTTACTACAGAACAAATGTGAT	907
OY	96	LEUTHRIHYSLSYVALVALASPTYRTHRIALALYSGLIUPHEALASPSERUEGLY	115
Db	908	ATGACGACCAAAAGTGTACTACCTAAAGAAAAGAGACAGCTTCAAGGAGCAATGCT	967
OY	116	ILIEPHELEUGLUTHSERALALYASNAALATHASVALGLUGLUSERPHEMETHR	135
Db	968	ATTAGGTGTTTGGACCTAGGTCACAAAGCAAAATTAATCAATCAAAAAGGCTTCTCCAG	1027
OY	136	METALALAGLIULEYLSYASATGMEGLYPROGLYALATHRIALAGLYALAGLUYS	155
Db	1028	TTTAGCTGAAGTATC	1042
OY	156	SERASNAVALYSLIEGLINSETRHPRVALYSGIN-----	167
Db	1043	-----CTTGAAAGACCCCTGTATTAAGAGCCCAACAGTAAATGTAGATATC	1090
OY	168	-----SergIyGLyGLY-----CysCys 173	

Mon Jun 23 16:54:08 2003

us-09-820-003b-2.p2n.rnpb

Page 10

Db 1091 ACAGGTGGAGGAGGCGCTGACAGGCTGGAAGAGCAATGCTGC 1132

Search completed: June 22, 2003, 19:48:03  
Job time : 154 secs







QY	90	ValGI <sup>1</sup> Asn <sup>1</sup> IL <sup>1</sup> SC <sup>1</sup> YS <sup>1</sup> Asp <sup>1</sup> Leu <sup>1</sup> Thr <sup>1</sup> Thr <sup>1</sup> ys <sup>1</sup> Val <sup>1</sup> Val <sup>1</sup> Asp <sup>1</sup> Y <sup>1</sup> Tr <sup>1</sup> Tr <sup>1</sup> Tr <sup>1</sup> Ala <sup>1</sup> ys <sup>1</sup> Iu	109
Db	420	GTGGGCAACAAGACGACCTCACACCAAGAGAGTGTGTGCAACACACAGCAAGAG	479
QY	110	Pha <sup>1</sup> IL <sup>1</sup> Asp <sup>1</sup> Ser <sup>1</sup> Ieu <sup>1</sup> GI <sup>1</sup> Y <sup>1</sup> Le <sup>1</sup> Pro <sup>1</sup> Ph <sup>1</sup> Ieu <sup>1</sup> GI <sup>1</sup> Y <sup>1</sup> Thr <sup>1</sup> Ser <sup>1</sup> Ala <sup>1</sup> ys <sup>1</sup> Asn <sup>1</sup> Ala <sup>1</sup> Thr <sup>1</sup> Asn <sup>1</sup> Al	129
Db	480	TTTGGAGACTCTCTGGGCAATCCCTCTTGTGGAGCGAGGCCCAAGAAATGGCAACCAATGTC	539
QY	130	GI <sup>1</sup> GI <sup>1</sup> IL <sup>1</sup> Ser <sup>1</sup> Pha <sup>1</sup> Met <sup>1</sup> Tr <sup>1</sup> Met <sup>1</sup> Ala <sup>1</sup> Ia <sup>1</sup> Iu <sup>1</sup> IL <sup>1</sup> ys <sup>1</sup> ys <sup>1</sup> ys <sup>1</sup> Asp <sup>1</sup> Met <sup>1</sup> GI <sup>1</sup> Y <sup>1</sup> Pro <sup>1</sup> GI <sup>1</sup> Y <sup>1</sup> Ala <sup>1</sup> Thr	149
Db	540	GACACAGCGTTCAGACCAATGCGCTGTCAATCAAAAACGGATGGGGCCTGTGGACACCC	599
QY	150	Ala <sup>1</sup> GI <sup>1</sup> GI <sup>1</sup> Y <sup>1</sup> Ala <sup>1</sup> GI <sup>1</sup> ys <sup>1</sup> Ser <sup>1</sup> Asn <sup>1</sup> Val <sup>1</sup> IL <sup>1</sup> GI <sup>1</sup> Ser <sup>1</sup> Tr <sup>1</sup> Pro <sup>1</sup> Val <sup>1</sup> IL <sup>1</sup> ys <sup>1</sup> GI <sup>1</sup> Asn <sup>1</sup> GI <sup>1</sup> Ser <sup>1</sup> Y <sup>1</sup>	169
Db	600	TC <sup>1</sup> TGGGGGCGC---GAGCGGCCCAATCTCAAGTACAGACGACCCTGTAAAGCGGGCTGGC	656
QY	170	GI <sup>1</sup> Y <sup>1</sup> GI <sup>1</sup> Y <sup>1</sup> GI <sup>1</sup> ys <sup>1</sup> ys <sup>1</sup>	173
Db	657	GGTGGCTGTGGC	668

```

US-09-399-913-66
; Sequence 66, Application US/09399913
Patent No. 6361971
; GENERAL INFORMATION:
Applicant: Rhodes, Kenneth
Applicant: Betty, Maria
Applicant: Ling, Hui-Ling
Applicant: An, Weiqiang
TITLE OF INVENTION: FORTRANISM CHANNEL INTERACTORS AND USES THEREFOR
FILE REFERENCE: NRI-07/0CP2
CURRENT APPLICATION NUMBER: US/09/399, 913
EARLIER FILING DATE: 1999-09-21
EARLIER APPLICATION NUMBER: USSN 60/110, 277
EARLIER FILING DATE: 1998-11-30
EARLIER APPLICATION NUMBER: USSN 60/110, 033
EARLIER FILING DATE: 1998-11-25
EARLIER APPLICATION NUMBER: USSN 60/109, 333
EARLIER FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: USSN 09/298, 731
EARLIER FILING DATE: 1999-04-23
EARLIER APPLICATION NUMBER: USSN 09/350, 614
EARLIER FILING DATE: 1999-07-09
EARLIER APPLICATION NUMBER: USSN 09/350, 874
EARLIER FILING DATE: 1999-07-09
NUMBER OF SEQ ID NOS: 73
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 66
LENGTH: 639
TYPE: DNA
ORGANISM: Rattus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(636)
US-09-399-913-66

```

Alignment Scores:	
Pred. No.:	2,52e-36
Score:	333.00
Percent Similarity:	52.86%
Best Local Similarity:	34.76%
Query Match:	37,588
DB:	4
Length:	633
Matches:	73
Conservative:	38
Mismatches:	55
Indels:	4
Gaps:	2

[illegible]

Oy	48	Phenylalanylthreonylleucineasparglycylserylthreonylleu	glnile-----	64
Db	127	TTTTGGCTCGCAATGAAACACATTGATGGGAACAGATAAACTCACA	ATCTGGATACA	186
Oy	64	-----		64
Db	187	GCAGGCGACGAGTCCCTCGTCTCATCAACAGTCATATTACAGAGT	CGCAGCGGGGCT	241
Oy	65	-----	-GluSerPheasnAsnValIysGlnI	75
Db	247	TTACTAGTGTATGATATTACAGAGACGACGTTCAACACTGACAC	CTGTATGAA	304
Oy	76	GluIleAspArgTyrAlaSerGluAsnValAsnIysIleuValGlyI	asnIysCysAsp	95
Db	307	GAGCCCGTCACATCCATTCACATCGGTCACTCATCTGTTATGGAA	TAAANCTGC	366
Oy	96	LeuThrThyIysValValAspTyrThrAlaIysGluPheAlaAsp	SerLeuGly	111
Db	367	TTGGAATCTAGAGAGAAAGTGAAGAAAAGGAAAGAGTGAAGCTT	TGGCACAGACATGGA	422
Oy	116	IleProPheLeuGlnThrSerAlaIysAsnAlaPheAsnValGlnI	serPheMetThr	133
Db	427	CTTATCTCTATGAAACTCTGCCAAGACTGCTCTAATGTASAGAG	GGCTTTATTAC	488
Oy	136	MetAlaIleGluIleIysIysArgMetGlyProGlyAlaThrAla	IysAlaGluIys	155
Db	487	ACAGCAAAAGAAATTTATGAAATAATCCACAGAGAGGGCTCTT	GGACATTAAATATGAGCA	546
Oy	156	SerAsnValIysIleGlnSerThrProVal-----		165
Db	547	AACGGCATCAAAATATGGCCCTCAGCATGCTCTCAACCATGAT	CTCACAGGACCAACNA	606
Oy	166	-----	LysGlnSerGlyIysGlyIysCys	173
Db	607	GGAGGGACACAGCGACGGGGAGAGGCTGCTCC		636

RESULT 4  
US-08-888-077A-28  
Sequence 28, Application US/08888077A  
Patent No. 6020143  
GENERAL INFORMATION:  
APPLICANT: ST. GEORGE-HYSLOP, PETER H  
APPLICANT: ROMMENS, JOHANNA M  
APPLICANT: FRASER, PAUL E  
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USBS THEREFOR  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUWOLZ & MENTLIK  
STREET: 600 SOUTH AVENUE WEST  
CITY: WESTFIELD  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/888,077A  
FILING DATE: 03-JUL-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/592,541  
FILING DATE: 26-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: PALISI, THOMAS M  
REGISTRATION NUMBER: 36,629  
REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP CIP IV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 654-5000



Oy	65	-----GlutserPheaaNaenVallys	71
Db	303	GCTCAGGCGCTTCGTGCTATGCATACCAACAATACTTCTTGACAAACATCAGG	365
Oy	72	GLTPRLNGINGLIILLeAspargylrLleuSerGIuaNvalAanlyLleuIeuValglys	91
Db	363	GCCCTGGCTACTACGAGATTCAATAGATRTCCAGAGGAGCATGTGATACCTGCTAGGC	422
Oy	92	AenlyAcysAspleuthrThrlystysValValAsptyrThrThralAllysGIuePhala	111
Db	423	AACACGCGGATATAGACACGAAGAAGATGTACTCGTCCGAAGACGAGACCTTGACC	482
Oy	112	AsperleuglyLlePropheluengLuthrSeralAlaysnaIarhaNvalGIuin	131
Db	483	AGGAGACTAGGGGTGTCTTCCTTCCTCGAGACGCCCAACACTGCATGATGTGGAGTTA	542
Oy	132	SerPheMetThrMetaIalagLluelLylysArgMecgly	145
Db	543	GCCTTCTGGCCATCGCAAGAACACTGAATAACGGGCGCGG	584
<b>RESULT 6</b>			
US-09-484-970B-142			
Sequence 142, Application US/09484970B			
Patent No. 6426186			
<b>GENERAL INFORMATION:</b>			
APPLICANT: Jones, Karen A.			
APPLICANT: Volkmutz, Rayne A.			
APPLICANT: Walker, Michael G.			
TITLE OF INVENTION: BONE REMODELING GENES			
FILE REFERENCE: PB-0014 US			
CURRENT APPLICATION NUMBER: US/09/484,970B			
CURRENT FILING DATE: 2000-01-18			
NUMBER OF SEQ ID NOS: 172			
SOFTWARE: PERL Program			
SEQ ID NO 142			
LENGTH: 2612			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: misc.feature			
OTHER INFORMATION: Incyte ID No. 6426186 412477.ICB1			
US-09-484-970B-142			
<b>Alignment Scores:</b>			
Pred. No.: 7		35e-33	Length: 2612
Score: 314.50			Matches: 61
Percent Similarity: 56.90%			Conservative: 36
Best Local Similarity: 35.06%			Mismatches: 42
Query Match: 35.50%			Indels: 33
Db: 4		Gaps: 2	
US-09-820-003B-2 (1-173) x US-09-484-970B-142 (1-2612)			
Oy	5	AsnProglutyrAspyrLleuPheLysIleuIeuIlleglyAspseryValglylys	24
Db	82	AGTCCAGCTTAGACACCTCACGGGCAAGGTGATGCTCTGTGGAGACACAGCGCTGGCAA	141
Oy	25	SerCyLeuIleuLleuLargPheAlasparThyr---ThrgLuserYrLlleSerThr	43
Db	142	ACAATGTTCTTCATGCCAATTCAAAGACGGGCGCTTCGTCCGAGACCTTCATAGCCAC	201
Oy	44	IleGlyValAsPheLysIleArgThrIleuIeuAspClYstHrIleYsIeuIn	63
Db	202	GTCGGCATAGACTTACGAACAAAGGTGTACTGTGATGCGGTGAGAAAGTAAGCTCGAG	261
Oy	64	Ile-----	64
Db	262	ATCTGGGACACCGCTGGGACGAGACGGTTCGAAAGCGTACCCATGCTTATTTACAGAGAT	321
Oy	65	-----GluSerPheaaNaenVallys	71
Db	322	GCTCAGGCGCTTCGTGCTATGCATACCAACAATACTTCTTGACAAACATCAGG	381

```

OY 7 GIATPLeuGlnIuIuIAspATgYrYAlAscRlAsnVAlAsnYsLeuValGly 91
Db 382 GCCTGCTCAGTACAGATTCATGAGTATGCCAGAGGGAGCTGTGATCATGCTGCTAGG 444
OY 92 AsnYsCysAspLeuThrThyLysValValAspYrThrThrAlaLysGluPheAla 111
Db 442 AACAGAGCGGATATGACGAGGAAAGATGATCGTTCCGAAAGACGAGAGACTTGCC 560
OY 112 AspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaIrnAsValGluGln 133
Db 502 AGGAGATrAGGCTTCTCCCTCTGAGACCGACCGCAGACTGCGATGATGTGAGATrA 560
OY 132 SerPheMetThrMetAlaAlaGlnIleLysLysArgMetCyl 145
Db 562 GCCCTTCTGCGCTCCAGCAAGAACTGAAATACCGGAGCGG 603

RESULT 7
; Sequence 2, Application US/08824873
; Patent No. 5843717
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl
; TITLE OF INVENTION: NOVEL RAB PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PASTESO for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824, 873
; FILING DATE: Filed Hewlett
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0240 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1340 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PANCNOT04
; CLONE: 738957
; US-08-824-873-2

Alignment Scores:
Prod. No.: 6,31e-29 Length: 1340
Score: 282.50 Matches: 70
Percent Similarity: 51.208 Conservative: 37
Best Local Similarity: 33.498 Mismatches: 45
Query Match: 31.888 Indels: 57
DB: 2 Gaps: 5

US-09-820-003b-2 (1-173) x US-08-824-873-2 (1-1340)
8 TyrAspYrIleuPheLysLeuLeuLeuIleGlyAspSerGlyValIleYrSerCysLeu 27

```

```

|||||
Db 22 TACGACGTGCGCTTACGATGTCACACAGAGCGCTTGTGCAACATCCAGCGCTGCTG 81
QY 28 Leu-LeuArgPheAlaAspAspThrTyr---ThrGluSerTyrIleSerThrIleGlyVa 46
Db 82 CTGGGTGGGATTCAGAGATGTCCTTCTGCGGGGAGCCTTCATCTCCACCGTAGC-AT 140
QY 46 AspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLysLeuGlnIle---- 64
Db 141 TGACTTCCGGAACAAAGTTGCGAGCTGATGTGTGAAGTGAAGCTGCAGATGTGGGA 200
QY 64 ----- 64
Db 201 CACAGCTGTGACGAGCGGTTCCGAGTGTACCAATGCTACTACCGGGATGCTCAGTC 260
QY 65 -----GluSerPheAsnAsnValLysGlnTyrLe 74
Db 261 TCTGCTGCTGCTTACGATGTCACACAGAGCGCTTGTGCAACATCCAGCGCTGCT 320
QY 74 uGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeuValGlyAsnLysCy 94
Db 321 GACCGAGATCCAGAGTACGCCACAGCAGCTGCGCTCATGCTGCGGGAACAAGGT 380
QY 94 sAspLeuThrThrLysLysValValAspTyrThrThrAlaLysGluPheAlaAspSerIle 114
Db 381 GGACTCTGCCATGAGCGCTGTGTGAAGAGGAGAGCGGGGAGAGCTGGCCCAAGAGTA 440
QY 114 uGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnValGlnGlnSerPheMe 134
Db 441 TGGACTGGCCCTTCATGAGAGACAGCGCCCAAGCGGCGCTCAACGTCGACTTGCCTTAC 500
QY 134 ThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThrAlaGlyAlaG1 154
Db 501 AGCCATACCAAGAGTTGAGACAGCGC----- 528
QY 154 uLysSerAsnValLysIleGlnSerThrPro-----ValLysG1 167
Db 529 -----TCCATGAAGGCTCCACAGAGCGCGCTTCCGGCTGCATGATTACGTTAAGAG 581
QY 167 nSerGlyGlyGly-----CysCys 173
Db 582 GGAGGTCGAGGGGCGCTCTGCTGC 606

```

## RESULT 8

```

US-09-198-184-2
; Sequence 2, Application US/09198184
; Patent No. 6010859
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL RAB PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,184
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/824,873
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.

```

```

; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0240 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1340 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PANCNOT04
; CLONE: 738957
; US-09-198-184-2

```

## Alignment Scores:

Pred. No.:	6,31e-29	Length:	1340
Score:	282.50	Matches:	70
Percent Similarity:	51.20%	Conservative:	37
Best Local Similarity:	33.49%	Mismatches:	46
Query Match:	31.88%	Indels:	57
DB:	3	Gaps:	5

US-09-820-003b-2 (1-173) x US-09-198-184-2 (1-1340)

```

QY 8 TyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyValGlyLysSerCysLeu 27
Db 22 TACGACGTGCGCTTACGATGTCACACAGAGCGCTTGTGCAACATCCAGCGCTGCTG 81
QY 28 Leu-LeuArgPheAlaAspAspThrTyr---ThrGluSerTyrIleSerThrIleGlyVa 46
Db 82 CTGGGTGGGATTCAGAGATGTCCTTCTGCGGGGAGCCTTCATCTCCACCGTAGC-AT 140
QY 46 AspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLysLeuGlnIle---- 64
Db 141 TGACTTCCGGAACAAAGTTGCGAGCTGATGTGTGAAGTGAAGCTGCAGATGTGGGA 200
QY 141 ----- 64
Db 201 CACAGCTGTGACGAGCGGTTCCGAGTGTACCAATGCTACTACCGGGATGCTCAGTC 260
QY 65 -----GluSerPheAsnAsnValLysGlnTyrLe 74
Db 261 TCTGCTGCTGCTTACGATGTCACACAGAGCGCTTGTGCAACATCCAGCGCTGCT 320
QY 74 uGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeuValGlyAsnLysCy 94
Db 321 GACCGAGATCCAGAGTACGCCACAGCAGCTGCGCTCATGCTGCTGGGGAACAAGGT 380
QY 94 sAspLeuThrThrLysLysValValAspTyrThrThrAlaLysGluPheAlaAspSerIle 114
Db 381 GGACTCTGCCATGAGCGCTGTGTGAAGAGGAGAGCGGGGAGAGCTGGCCCAAGAGTA 440
QY 114 uGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnValGlnGlnSerPheMe 134
Db 441 TGGACTGGCCCTTCATGAGAGACAGCGCCCAAGCGGCGCTCAACGTCGACTTGCCTTAC 500
QY 134 ThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThrAlaGlyAlaG1 154
Db 501 AGCCATACCAAGAGTTGAGACAGCGC----- 528
QY 154 uLysSerAsnValLysIleGlnSerThrPro-----ValLysG1 167
Db 529 -----TCCATGAAGGCTCCACAGAGCGCGCTTCCGGCTGCATGATTACGTTAAGAG 581
QY 167 nSerGlyGlyGly-----CysCys 173
Db 582 GGAGGTCGAGGGGCGCTCTGCTGC 606

```

## RESULT 9

```

US-08-773-423-4
; Sequence 4, Application US/08773423

```

Patent No. 5869291  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
APPLICANT: Bandman, Olga  
TITLE OF INVENTION: NOVEL RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/773,423  
FILING DATE: Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0183 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 847 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Consensus  
CLONE: Consensus  
US-08-773-423-4  
Alignment Scores:  
Pred. No.: 1.81e-25 Length: 847  
Score: 255.00 Matches: 66  
Percent Similarity: 49.538 Conservative: 39  
Best Local Similarity: 31.134 Mismatches: 53  
Query Match: 28.784 Gaps: 54  
DB: 2 Gaps: 5  
US-09-820-003b-2 (1-173) x US-08-773-423-4 (1-847)  
QY 7 GIUTYRASPYYRIleuPhelyleuileGlyAspserglyValglyserCys 26  
DB 92 GATATATACCTTGTCTCAGGTGTCTGATGGGCAATGATGGGGAAGCAAT 151  
QY 27 LeuLeuLeuArpPhelAspserPheThyThrGlyserTyrlleSerThrleGlyAla 46  
DB 152 CTACCTCCCGATTCCAGCGCAATGATGAGCCACAGACCCCAACCATCGGGTT 211  
QY 47 AspPheyleuArpPhelThrleGlyLeuAspGlyLysThrleLysLeuGlnle----- 64  
DB 212 GAGTTTCACACCGCACTGTATGTGTGGGACCGCTGCTGCAAGGCTCGATCTGGGAC 271  
QY 64 ----- 64  
DB 272 ACAGCTGCGTGGAGCGGTACCGACCATCCCTGGCGACTACTGCTGCTGACGTGGGG 331  
QY 65 -----GluSerPheAsnAsnValLysGlnThrPleu 74  
:::|||||

DB 332 GCCCTCTGCTGTTTACCTAACACAGACACCTATGCTGTGTGGAGCATGCTGC 391  
QY 75 GlnLILIASPARGYTRAlAserGluAsnValAsnLysleuLeuValGlyAsnLysCys 94  
DB 392 AAGACGCTCATGACACATGCTGAGAGCCAGATCGCTGCTGCTGCTGCTGCTGCTGCT 451  
QY 95 AspLeuThrThrLysLysValAlaAspTYRThrAlaLysGluPheAlaAspSerLeu 114  
DB 452 GACCTCAGCAGGCGGAGAGTGGCCACTGAGAGGCGCCAGATGTGCTGAAACAT 511  
QY 115 GlyLeuPheLeuGluThrSerAlaLysAsnAlaThrAsnValGluGlnSerPheMet 134  
DB 512 GAGCTGCTCTCTGAGACCTGACGCTGAGCTTACCAATGTGACTGACTTGGAG 571  
QY 135 ThrMetAlaLglnle-----LysLysArgMetGlyPro----- 146  
DB 572 ACTGCTGTAAGAAATTTTGGAAAGTGTCACAGACAGACAGACATCCGAGCC 631  
QY 147 ---GlyAlaThrAlaGlyLysAlaGlyLysSerAsnValLysIleGlnSerThrProVal 165  
DB 632 AATGCCATCATCTGCGAGGCGCCAGGNTGGA----- 664  
QY 166 LysGlnSerGlyGlyLys-----CysCys 173  
DB 665 CAGAGCGCTGCCCTGGGAGAGAGAGGCGCTGTTC 700  
RESULT 10  
US-08-741-411-2  
Sequence 2, Application US/08741411  
Patent No. 6124116  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Au-Yang, Janice  
TITLE OF INVENTION: NOVEL RAB PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/741,411  
FILING DATE: Herewith  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0139 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 848 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE: Consensus  
US-08-741-411-2

Alignment Scores:

Pred. No.:	2,966-22	Length:	848
Score:	231.50	Matches:	53
Percent Similarity:	45.45%	Conservative:	32
Best Local Similarity:	28.34%	Mismatches:	65
Query Match:	26.13%	Indels:	37
DB:	3	Gaps:	3

US-09-820-003b-2 (1-173) x US-08-741-411-2 (1-848)

```

QY 19 AspSerGlyValGlyLysSerCysLeuLeuAlaGlyPheAlaAspAspThrThrGlu 38
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 216 GACACTGGGGTGGGAATCAAGCATCGTGTGATGATTCAGAGATCATTGGACAC 275
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 39 SerThrIleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLys 58
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 276 AACATCAGCCCTCATCTGGGCGCATCTTTATGACCAAAACGTGCTGTGGAATGAA 335
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 ThrIleLysLeuGlnIle----- 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 336 CTTCACAAGTTCCTCATCTGGGACACTGCTGTCAGAACGGTTTCATTCATTGGCTCC 395
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 65 -----GluSer 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 336 ATGTACTTCAGGCTCAGCTGACCTGTTATCGTGTATGATATTCACAGAGATTGA 455
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 67 PheAsnAsnValLysGlnThrPheGlnGlnIleAspArgTyraLysSerGluAsnValAsn 86
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 456 TTTTATACCTTGAAATAATGGGTCAAGAGCTGAAAGAACATGATGCCAATAACATGTA 515
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 87 LysLeuLeuValGlyAsnLysCysAspLeuThrThrLysValValAspTyThrThr 106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 516 ATGCCATCGCTGGAAACAGTGCAGCTTCAGATATATGAGAGAGTCCCTGAAGAT 575
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 107 AlalysGluPheAlaAspSerLeuGlyLeuPheGluThrSerAlalysAsnAla 126
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 576 GCTTAAAGATACGCTGAATCCTAGTGCATGCTGTGGAACAAATGCAAAATGCT 635
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 127 ThrAsnValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyPro 146
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 636 ATTATATCGAAGACCTCTTTCMA-----GGATACAGCCGCCAGATCCACACC 683
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 147 GlyAlaThrAlaGlyLysIleGlyLysSerAsnValLysIleGlnSerThrProValLys 166
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 684 TTGACCCCATGAAATGGAACATGGAACATCAATCAAGTGTGAGAAG---CCAACCAATG 740
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 167 GlnSerGlyGlyGlyCysCys 173
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 741 CAAGCCAGCCCGCGGTCTGT 761
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 11  
US-08-429-964-85  
Sequence 85, Application US/08429964  
Patent No. 5962243  
GENERAL INFORMATION:  
APPLICANT: BROWN, MICHAEL S.  
APPLICANT: GOLDSTEIN, JOSEPH L.  
APPLICANT: REISS, YUVAL  
APPLICANT: JAMES, GUY L.  
TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL  
NUMBER OF SEQUENCES: 85  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE & DUKKEE  
STREET: P.O. BOX 4433  
CITY: HOUSTON  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,964  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,625  
FILING DATE: 16-FEB-1993  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 07/822,011  
FILING DATE: ABANDONED  
CLASSIFICATION: 435  
APPLICATION NUMBER: PCT/US/91/02650  
FILING DATE: 18-APR-1991  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 07/615,715  
FILING DATE: 20-NOV-1990  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 07/510,706  
FILING DATE: 18-APR-1990 (ABANDONED)  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: PARKER, DAVID L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UTSD:432/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 85:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 607 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-429-964-85

Alignment Scores:

Pred. No.:	4,146-21	Length:	607
Score:	221.50	Matches:	55
Percent Similarity:	49.72%	Conservative:	35
Best Local Similarity:	30.39%	Mismatches:	56
Query Match:	25.00%	Indels:	35
DB:	2	Gaps:	4

US-09-820-003b-2 (1-173) x US-08-429-964-85 (1-607)

```

QY 12 PheLysLeuLeuLeuIleGlyAspSerGlyValGlyLysSerCysLeuLeuAlaGlyPhe 31
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 10 TATAAAGCTTGGTGGTAGTGGACCTTGTGGCGTGAAGAGCTTGCATACGATACAGCTA 69
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 32 AlaAspAspThrThrThrGluSerThrIleSerThrIleGlyValAspPheLysIleArg 51
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 70 ATTGACATCATTTTGTGACCAATATGATCCACAAATAGAGATCTCTACAG---AAG 126
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 52 ThrIleGluLeuLeuAspGlyLysThrIleLysLeuGlnIle----- 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 127 CAAGTAGTAATGATGAGAAACCTGCTCTTGATATTCGACACAGAGGTCAAGAG 186
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 64 ----- 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 187 GAGTACAGTGCATAGGAGGACAGTACATGAGAGCTGGGAGGCGCTTCTTGTGTATTT 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 65 -----GluSerPheAsnAsnValLysGlnThrPheGlnGlnIleAspArg 79
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 247 GCCATTAATATATCTTAATCATTTGAAGATATTCACCTTTATGAGAACAAATTAAGA 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 80 TyraLal-----SerGluAsnValAsnLysLeuLeuValGlyAsnLysCysAspLeuThrThr 98
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 307 GTTAAAGACTCGAAGAGTACCTTGTGCTCCTAGAGAAATATGATTTGCTCTCT 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 99 LysLysValValAspTyThrThrAlalysGluPheAlaAspSerLeuGlyLeuPhe 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

[illegible]





```

0y      TyrIala-----SerGIuaanValaSenYLSLeuauValaGlyasnLYsCysAspLeuThr 98
Db      307 GTTAAAGACTCTGAAGAATGTACCTTAGTGCTTACTAGTAAGAAATAATGATTGCCTTCT 366
0y      LysIsvaValaIasPYrTrThrAlaYsgLuphealAaspSerLeuGLYLleProPhe 118
Db      367 ---AGAACGTGTGACACAAAACAAGCGTCAGAGCACTTAGCAAGAGTTATGGAATTCCTTT 423
0y      LeuGIuThrSeRAlaLySaSnAlaThrsAnVaGlucInserPIemEtThrMetIala 138
Db      424 ATTCAAACATCATGCCAAAGACAAAGACAGACAGACTGGAGATGCTTTTATATCATGTGGAGAA 483
0y      139 GluILeUgLyIs 142
Db      484 GAGNTCCGACAA 495

RESULT 15
US-09-075-454-13
Sequence 13, Application US/09075454
Patent No. 6391580
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Tang, Y. Tom
APPLICANT: Lal, Preeti
APPLICANT: Guglier, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Patterson, Chandira
APPLICANT: Batra, Sajeev
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: RAS PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075,454
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/766,551
FILING DATE: DECEMBER 12, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Certone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 803 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: OVARUT10
CLONE: 2703745
US-09-075-454-13

Alignment Scores:
Pred. No.:          2,24e-20           Length:      803
Score:              217.50             Matches:     60
Percent Similarity: 46.88%             Conservative: 45

```

Best Local Similarity:	26.79%	Mismatches:	61
Query Match:	24.55%	Indels:	58
DB:	4	Gaps:	5

US-09-820-003B-2 (1-173) x US-09-075-454-13 (1-803)

QY 3 SerMetAsnProGluTyrAspTyrLeu----- 11

Db 3 GCACCTCTCGCCCCAGACAGACCTGGTAGATGACAAACGCTAAATCTTTAGCTCACAGAAG 62

Lz ---phenylsLeuleuleuIleglyAspserglyValglyLyssercysleuleuleu-Ar 30

DD 05 GCLACAGAAIGIACIIICGCGGACGCIGCAGTGGGGAAGICTAGTTCCCAATGGAG 12

[illegible][illegible][illegible]

Curriculum

[illegible]

OY 65 -----GluSerPheAsnValIvsGIntRpleGIngIuIleAs 78

Db 303 GTATGATGTACATGTGAGAAAGCTTTCTTAACATACGAGATGGGTAGATATGATTGA 36

78 parGTyRAIaSerGIuaSnValaSnLysLeuLeuValGIyASnLysCysaspLeuThrTh 98

Db 363 GGATGCAGCCCATGAGACTGTCCCATTAATGCTGGTAGGAAACAAGCGTGACATTCGTGA 42

QY 98 L-----LysLysValValAspTyrThrThrAlaLysGluPheAlaAs 11

Db 423 CACTGCTGCTACAGAGGACAAAAATGTGTCCAGGGCACTTTGGAGAGAAACTGGCCAT 48

112 pserLeuGlyIleProPheLeuGluThrserAlaLysAsnAlaThrAsnValGluGlnSe 13

403 GATCGTATGGGGCTATATCTCTGTGTTTCTTATGCTTATGCTGATGGCT 34

```

13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87 89 91 93 95 97 99 101 103 105 107 109 111 113 115 117 119 121 123 125 127 129 131 133 135 137 139 141 143 145 147 149 151 153 155 157 159 161 163 165 167 169 171 173 175 177 179 181 183 185 187 189 191 193 195 197 199 201 203 205 207 209 211 213 215 217 219 221 223 225 227 229 231 233 235 237 239 241 243 245 247 249 251 253 255 257 259 261 263 265 267 269 271 273 275 277 279 281 283 285 287 289 291 293 295 297 299 301 303 305 307 309 311 313 315 317 319 321 323 325 327 329 331 333 335 337 339 341 343 345 347 349 351 353 355 357 359 361 363 365 367 369 371 373 375 377 379 381 383 385 387 389 391 393 395 397 399 401 403 405 407 409 411 413 415 417 419 421 423 425 427 429 431 433 435 437 439 441 443 445 447 449 451 453 455 457 459 461 463 465 467 469 471 473 475 477 479 481 483 485 487 489 491 493 495 497 499 501 503 505 507 509 511 513 515 517 519 521 523 525 527 529 531 533 535 537 539 541 543 545 547 549 551 553 555 557 559 561 563 565 567 569 571 573 575 577 579 581 583 585 587 589 591 593 595 597 599 601 603 605 607 609 611 613 615 617 619 621 623 625 627 629 631 633 635 637 639 641 643 645 647 649 651 653 655 657 659 661 663 665 667 669 671 673 675 677 679 681 683 685 687 689 691 693 695 697 699 701 703 705 707 709 711 713 715 717 719 721 723 725 727 729 731 733 735 737 739 741 743 745 747 749 751 753 755 757 759 761 763 765 767 769 771 773 775 777 779 781 783 785 787 789 791 793 795 797 799 801 803 805 807 809 811 813 815 817 819 821 823 825 827 829 831 833 835 837 839 841 843 845 847 849 851 853 855 857 859 861 863 865 867 869 871 873 875 877 879 881 883 885 887 889 891 893 895 897 899 901 903 905 907 909 911 913 915 917 919 921 923 925 927 929 931 933 935 937 939 941 943 945 947 949 951 953 955 957 959 961 963 965 967 969 971 973 975 977 979 981 983 985 987 989 991 993 995 997 999 1001 1003 1005 1007 1009 1011 1013 1015 1017 1019 1021 1023 1025 1027 1029 1031 1033 1035 1037 1039 1041 1043 1045 1047 1049 1051 1053 1055 1057 1059 1061 1063 1065 1067 1069 1071 1073 1075 1077 1079 1081 1083 1085 1087 1089 1091 1093 1095 1097 1099 1101 1103 1105 1107 1109 1111 1113 1115 1117 1119 1121 1123 1125 1127 1129 1131 1133 1135 1137 1139 1141 1143 1145 1147 1149 1151 1153 1155 1157 1159 1161 1163 1165 1167 1169 1171 1173 1175 1177 1179 1181 1183 1185 1187 1189 1191 1193 1195 1197 1199 1201 1203 1205 1207 1209 1211 1213 1215 1217 1219 1221 1223 1225 1227 1229 1231 1233 1235 1237 1239 1241 1243 1245 1247 1249 1251 1253 1255 1257 1259 1261 1263 1265 1267 1269 1271 1273 1275 1277 1279 1281 1283 1285 1287 1289 1291 1293 1295 1297 1299 1301 1303 1305 1307 1309 1311 1313 1315 1317 1319 1321 1323 1325 1327 1329 1331 1333 1335 1337 1339 1341 1343 1345 1347 1349 1351 1353 1355 1357 1359 1361 1363 1365 1367 1369 1371 1373 1375 1377 1379 1381 1383 1385 1387 1389 1391 1393 1395 1397 1399 1401 1403 1405 1407 1409 1411 1413 1415 1417 1419 1421 1423 1425 1427 1429 1431 1433 1435 1437 1439 1441 1443 1445 1447 1449 1451 1453 1455 1457 1459 1461 1463 1465 1467 1469 1471 1473 1475 1477 1479 1481 1483 1485 1487 1489 1491 1493 1495 1497 1499 1501 1503 1505 1507 1509 1511 1513 1515 1517 1519 1521 1523 1525 1527 1529 1531 1533 1535 1537 1539 1541 1543 1545 1547 1549 1551 1553 1555 1557 1559 1561 1563 1565 1567 1569 1571 1573 1575 1577 1579 1581 1583 1585 1587 1589 1591 1593 1595 1597 1599 1601 1603 1605 1607 1609 1611 1613 1615 1617 1619 1621 1623 1625 1627 1629 1631 1633 1635 1637 1639 1641 1643 1645 1647 1649 1651 1653 1655 1657 1659 1661 1663 1665 1667 1669 1671 1673 1675 1677 1679 1681 1683 1685 1687 1689 1691 1693 1695 1697 1699 1701 1703 1705 1707 1709 1711 1713 1715 1717 1719 1721 1723 1725 1727 1729 1731 1733 1735 1737 1739 1741 1743 1745 1747 1749 1751 1753 1755 1757 1759 1761 1763 1765 1767 1769 1771 1773 1775 1777 1779 1781 1783 1785 1787 1789 1791 1793 1795 1797 1799 1801 1803 1805 1807 1809 1811 1813 1815 1817 1819 1821 1823 1825 1827 1829 1831 1833 1835 1837 1839 1841 1843 1845 1847 1849 1851 1853 1855 1857 1859 1861 1863 1
```

[illegible]

1000

170 [2] 173

Db 648 GAATTGTTGC 657  
|||||

```

search completed:  June 22, 2003, 18:30:30
Job time : 70 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 22, 2003, 17:34:10 ; Search time 1419 Seconds

(without alignments)  
1974.503 Million cell updates/sec

Title: US-09-820-003b-2  
Perfect score: 86  
Sequence: 1 MSSMNPEDYLFKLLIGDS.....EKSNWIKSTPVKSGSGGCC 173

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame.p2n.model -DEV-xip  
-O/cgnt2.1/USPFO.spool/US09820003/unalc\_18062003\_144130\_13104/app\_query.fasta\_1.327  
-DB-EST -OPMT-fastcap -SUFFIX-p2n.rst -MINMATCH-0.1 -LOOPEXT-0 -LOOPEXT-0  
-UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45  
-DOCFALGN-200 -THR SCORE-pct -THR MAX-100 -THR MIN-0 -ALIGN-15 -MODE-LOCAL  
-OUTFMT-pio -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000  
-USRR-US09820003 -CGN 1.1.2874 -gtnat\_18062003\_144130\_13104 -NCPV-6 -ICPV-3  
-NO\_MMAP -LARGEDECRY -NEG\_SCORES-0 -WAIT -DSPLOCK-100 -LONELOG  
-DEV\_TIMEOUT-120 -MAIN\_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6  
-FGAPEXT-7 -VGAPOP-10 -VGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database :  
EST:  
1: em\_estha.\*  
2: em\_esthm.\*  
3: em\_estln.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_estl.\*  
10: gb\_estl2.\*  
11: gb\_hic.\*  
12: gb\_estl3.\*  
13: gb\_estl4.\*  
14: gb\_estl5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_dln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_hum.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_fod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	860	97.1	948	AL539022	AL539022 AL539022
2	860	97.1	976	BM919860	BM919860 AGENCOURT
3	849	95.8	1130	BM542833	BM542833 AGENCOURT
4	847	95.6	862	BM223273	BM223273 AGENCOURT
5	847	95.6	944	BM926234	BM926234 AGENCOURT
6	831.5	93.8	646	BM093078	BM093078 AGENCOURT
7	825	93.1	957	BM219698	BM219698 AGENCOURT
8	822	92.8	864	BM437146	BM437146 AGENCOURT
9	820	92.6	901	BM944368	BM944368 AGENCOURT
10	819	92.4	1105	BM452262	BM452262 AGENCOURT
11	814	91.9	690	BM954378	BM954378 AGENCOURT
12	809	91.3	945	AL530265	AL530265 AL530265
13	808.5	91.3	642	BM073109	BM073109 AGENCOURT
14	807	91.0	872	BM684554	BM684554 AGENCOURT
15	806	91.0	855	BM739625	BM739625 AGENCOURT
16	801	90.4	807	BM698150	BM698150 AGENCOURT
17	800	90.3	696	BM786746	BM786746 AGENCOURT
18	800	90.3	890	BM437514	BM437514 AGENCOURT
19	795	89.7	673	BM692590	BM692590 AGENCOURT
20	795	89.7	846	BM619962	BM619962 AGENCOURT
21	794	89.6	1067	BM922791	BM922791 AGENCOURT
22	790	89.2	909	BM743352	BM743352 AGENCOURT
23	788.5	89.0	859	BM159480	BM159480 AGENCOURT
24	788.5	89.0	1034	BM067968	BM067968 AGENCOURT
25	788	88.9	611	AL1048520	AL1048520 hb61f06.r
26	787	88.8	828	BM1652154	BM1652154 AGENCOURT
27	783	88.4	913	BM158729	BM158729 AGENCOURT
28	783	88.4	913	BM970841	BM970841 AGENCOURT
29	780	88.0	933	BM068107	BM068107 AGENCOURT
30	778.5	87.9	588	BM074071	BM074071 AGENCOURT
31	777	87.7	901	BM450994	BM450994 AGENCOURT
32	777	87.7	937	BM1738727	BM1738727 AGENCOURT
33	776	87.6	676	BM913399	BM913399 AGENCOURT
34	775.5	87.5	612	BM9595367	BM9595367 AGENCOURT
35	775	87.5	1077	BM476507	BM476507 AGENCOURT
36	769.5	86.9	881	BM0424046	BM0424046 AGENCOURT
37	769	86.8	1047	BM800733	BM800733 AGENCOURT
38	768.5	86.7	643	AL535937	AL535937 AL535937
39	768.5	86.7	744	BM045935	BM045935 AGENCOURT
40	768.5	86.7	766	BM826700	BM826700 AGENCOURT
41	768.5	86.7	821	BM175331	BM175331 AGENCOURT
42	768.5	86.7	825	AL542911	AL542911 AL542911
43	768.5	86.7	826	BM561688	BM561688 AGENCOURT
44	768.5	86.7	837	AL520855	AL520855 AGENCOURT
45	768.5	86.7	862	BM941767	BM941767 AGENCOURT

## ALIGNMENTS

RESULT 1	AL539022	948 bp	MRNA	linear	EST 16-FEB-2001
LOCUS	AL539022	LTI FL013.Fbrn1	Homo sapiens	cDNA clone	COSDP030706 5
DEFINITION	AL539022	Prime, mRNA sequence.			
ACCESSION	AL539022	GI:12867866			
VERSION	AL539022.1	GI:12867866			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.				
	Full-length cDNA libraries and normalization				

## JOURNAL

Unpublished (2001)

COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

Source Location/Qualifiers

1. 948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSODF030YG06"  
/dev\_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
/lab\_host="DH10B"  
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 277 a 187 c 239 g 244 t 1 others

## ALIGNMENT SCORES:

Pred. No.: 1.87e-94 Length: 948  
Score: 860.00 Matches: 173  
Percent Similarity: 84.39% Conservative: 0  
Best Local Similarity: 84.39% Mismatches: 0  
Query Match: 97.07% Indels: 32  
DB: 9 Gaps: 1

US-09-820-003b-2 (1-173) x AL539022 (1-948)

OY 1 MetSerMetAsnProGluTyrAspTyrLeuPheLeuLeuLeuIleGlyAspSer 20  
DB 166 ATGTCACAGCAATCCGAAATATGATATTATTCAGTACTCTGATGCGACATCA 225  
OY 21 GLYVALIGLYSSErCysLeuLeuLeuArgPheAlaAspPThrTyrThrGluSerTyr 40  
DB 226 GGGGTGGAAAGCTCTGCTCTCTTTCAGTTTCAGATGATACATACAGAAAGCTAC 285  
OY 41 IleserThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
DB 286 ATCAGCACAAATGGTGGATTTCAAAATAGAACTATAGATTAGACGGGAAACAAATG 345  
OY 61 LysLeuGlnIle----- 64  
DB 346 AAGCTTCAAAATATGGGACACAGCAGCAAGAAAGATTTCGAACAATCACCCTCAGTTAT 405  
OY 65 -----GluSerPheAsn 68  
DB 406 TACAGAGAGCCCATGCATCATCTGTTGATGATGACAGATCAGAGCTCTTCAT 465  
OY 69 AsnValLysGlnTyrPheGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88  
DB 466 AATGTTAAACAGTGGCTCAGGAAATAGATGCTTATGTCAGAGAAATGTCACAAATG 525  
OY 89 LeuValIGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108  
DB 526 TTGCTAGAGGAACAAATGTGATCTGCACCAAAAGAAAGTAGTAGTACACAAACAGCAAG 585  
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128  
DB 586 GAATTTGCTGATCCCTGGATTCCTGTTTGGAAACAGAGCTGTAAGAAATGCAACGAT 645  
OY 129 ValGluGlnSerPheMetThrMetAlaIleGluIleLysArgMetGlyProGlyAla 148  
DB 646 GTAGAACAGTCTTTCATATGACATGCGACCTGAGATTAAAAAGCAATGGGTCCCGGACGA 705

OY 149 ThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168  
DB 706 ACAGCTGGTGGTGGTGAAGAGTCAATGTTAAATTCAGACACATCCAGTCAAGACATCA 765  
OY 169 GLYGLYGLYCysCys 173  
DB 766 GTGGAGGTGCTGTC 780

## RESULT 2

BM919860

LOCUS BM919860 976 bp mRNA linear EST 12-MAR-2002  
DEFINITION AGENCOURT.6708189 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5749635  
5', mRNA sequence.

ACCESSION BM919860  
VERSION BM919860.1 GI:19370239  
KEYWORDS EST.

## SOURCE

ORGANISM human.

REFERENCE Homo sapiens  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1. (bases 1 to 976)  
NIH-MGC http://mhc.nci.nih.gov/.

## AUTHORS

TITLE

Unpublished (1999)

## JOURNAL

Contact: Robert Strausberg, Ph.D.

## COMMENT

Email: cgabbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12779 row: f column: 04  
High quality sequence stop: 757.

## FEATURES

Source Location/Qualifiers

1. 976  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5749635"  
/clone\_id="NIH\_MGC\_120"  
/lab\_host="DH10B"  
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb. Insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 025. Note: this is a NIH\_MGC Library."

## BASE COUNT

285 a 196 c 246 g 248 t 1 others

## ORIGIN

## ALIGNMENT SCORES:

Pred. No.: 1.95e-94 Length: 976  
Score: 860.00 Matches: 173  
Percent Similarity: 84.39% Conservative: 0  
Best Local Similarity: 84.39% Mismatches: 0  
Query Match: 97.07% Indels: 32  
DB: 14 Gaps: 1

US-09-820-003b-2 (1-173) x BM919860 (1-976)

OY 1 MetSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
DB 177 ATGTCACAGCAATCCGAAATATGATATTATTCAGTACTCTGATGCGACATCA 236  
OY 21 GLYVALIGLYSSErCysLeuLeuLeuArgPheAlaAspPThrTyrThrGluSerTyr 40  
DB 237 GGGGTGGAAAGCTCTGCTCTCTTTCAGTTTCAGATGATACATACAGAAAGCTAC 296  
OY 41 IleserThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60

DB 297 ATCAGACATTTGGTGGATTTCAAAATAGACTATAGAGTTAGACGGAAACATC 356  
61 Lysleuglnlle----- 64  
DB 357 AAGCTTCAAAATATAGGACACAGACAGCCAGAAAGATTTCACATCCTCTGATTAT 416  
65 -----Gluserpheasn 68  
DB 417 TACAGAGAGCCCGATCATCTAGTGTGTATGATGACAGATCAGAGATGACATTCAT 476  
69 AsnValIysGlnITrPLeuGlnIuIleAspArgTyrAlaSerGluAsnValAlaLysLeu 88  
477 AATGTTAAACAGTGGCTGACGAAATAGATCTATGCTGCAATGCAATGCAAAATG 536  
DB 89 LeuValIysAsnLysCysAspLeuThrThrLysValAlaAspTyrThrAlaLys 108  
537 TTGGTAGAGGAAACAAATGATCTGACACAAAGAAAGTAGAGCTACACACAGCGAG 596  
DB 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128  
597 GAATTCCTGATTCCTCGAATTCGTTTGGAAACCACTGCTAAGAAATGCAACGAT 656  
DB 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148  
657 GTAAACAGCTTTCATGACATGACAGCTGAGATTAAAAAGCATGGGTCCGAGACA 716  
DB 149 ThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168  
717 ACAGCTGGTGGCTGAGAGATCCAAATGTTAAATTCAGACATCCAGTCAAGACATCA 776  
DB 169 GlyGlyGlyCysCys 173  
777 GGTGAGGTGGCTGC 791  
DB 777 GGTGAGGTGGCTGC 791  
RESULT 3 1130 bp mRNA linear EST 20-FEB-2002  
BM542833  
LOCUS AGENCOURT\_6426035 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5520670  
DEFINITION 5', mRNA sequence.  
ACCESSION BM542833  
VERSION BM542833.1 GI:18772697  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strusberg, Ph.D.  
Email: rstrus@r-emall.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNL at:  
http://image.llnl.gov  
Plate: ILAM12185 row: a column: 23  
High quality sequence stop: 597.  
Location/Qualifiers  
1. 1130  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5520670"  
/clone\_11b="NIH\_MGC\_71"  
/tissue="type="lelomyosarcoma"  
/lab\_host="DH108 (phage-resistant)"  
/note="Organ: uterus; Vector: PCMV-SPOK6; Site: 1; Moti:  
Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dr.  
Average insert size 2.1 kb.

BASE COUNT 319 a 241 c 290 g 280 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.4e-93 Length: 1130  
Score: 849.00 Matches: 171  
Percent Similarity: 83.41% Conservative: 0  
Best Local Similarity: 83.41% Mismatches: 2  
Query Match: 95.82% Indels: 32  
DB: 13 Gaps: 1  
US-09-820-003b-2 (1-173) x BM542833 (1-1130)  
DB 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
178 ATGTCAGCATGAATCCCGAATATGATTTATTTATTCAGTACTCTCATGCGACCTCA 237  
DB 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspSerThrTyrThrGluSerTyr 40  
238 GGGGTGGAAAGTCTTCCTCTCTAGGTTTCACATGATACATATACAGAAAGCTAC 297  
DB 41 IleSerThrIleGlyValaAspPheLysIleArgTyrIleGluLeuAspGlyLysThrIle 60  
298 ATCAGCAATTTGGTGTGATTTCAAAATAGAACTATGAGTTAGACGGAAACAAATC 357  
DB 61 Lysleuglnlle----- 64  
358 AAGCTTCAAAATATGAGACACAGACAGCCAGAAAGATTTCGAACATCACCTCCAGTTAT 417  
DB 65 -----Gluserpheasn 68  
418 TACAGAGAGCCCATGACATCATAGTTGTATGATGTCAGATCAGAGATCTCTCATAT 477  
DB 69 AsnValIysGlnITrPLeuGlnIuIleAspArgTyrAlaSerGluAsnValAlaLysLeu 88  
478 AATGTTAAACAGTGGCTGACGAAATAGATCTATGCTGCAATGCAAAATGCAACAAATG 537  
DB 89 LeuValIysAsnLysCysAspLeuThrThrLysValAlaAspTyrThrAlaLys 108  
538 TTGGTAGAGGAAACAAATGATGATGACACAAAGAAAGTAGTACTACACACAGCGAG 597  
DB 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128  
598 GAATTCGATTCCTCGAATTCGTTTGGAAACCACTGCTAAGAAATGCAACGAT 657  
DB 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148  
658 GTGACACATCTTTCATGACATGACAGCTGAGATTAAAAAGCATGGGTCCGAGACA 717  
DB 149 ThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168  
718 ACAGCTGGTGGCTGAGAGATCCAAATGTTAAATTCAGACATCCAGTCAAGCATTC 777  
DB 169 GlyGlyGlyCysCys 173  
778 AGTGAAGTGTGCTGC 792  
DB 778 AGTGAAGTGTGCTGC 792  
RESULT 4  
LOCUS B0223273 862 bp mRNA linear EST 02-MAY-2002  
DEFINITION AGENCOURT\_7518264 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:606561  
ACCESSION B0223273  
VERSION B0223273.1 GI:20404673  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)



DB	ATATGGAGACACAGAGCCAGGAAAGATTTCGAACATTCACGTTATATACAGAGA	259
OY	65 -----GtserPheAsnValValys 71	
Db	260 GCCCATCGCATCATTAGTTGTGTGATGATGTACAGATCAGAGAGTCTTCATATACGTTAAA	319
OY	72 GATTTPLeuGlnGluIleAspArgPheLysSerGluAsnValAsnLysLeuValGly 91	
Db	320 CATTGGCTTCAGAGGAGATAGATGCTCAGCCAGAGGAAATGTCAACAGTTGTGGTAGGG	379
OY	92 AsnLysCysAspLeuThrThrLysLysValValAspPheThrThrAlaLysGluPheAla 111	
Db	380 AACCAATGTGACCTCGACCCCAAGAAAGGTGTGATCTACACACACAGCAAGGAATTGGCA	439
OY	112 AspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnValGluGln 131	
Db	440 GATTTCCTCTGGAAATTCATTTTGGAAACAGTGTCAAGAAACGACAGATGTGGAACAG	499
OY	132 SerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThrAlaGly 151	
Db	500 TCTTTATATACAGAGGAGCAGATTTAAACCGAAGGGTCTCTGACACTCAGCTGGGT	559
OY	152 G1YAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerG1YGlyGly 171	
Db	560 GGTCGCGAAGAGTCCATGTATTAATCCAGAGCAGCTCCAGTCAAGCAGTCAAGTGGAGGC	619
OY	172 CysCys 173	
Db	620 TCGCTGC 625	
RESULT 6		
BJ093078		
LOCUS		
DEFINITION	BJ093078 N1B Mochii normalized Xenopus early gastrula library	
ACCESSION	Xenopus laevis cDNA clone X141118.5, mRNA sequence.	
VERSION	BJ093078	
KEYWORDS	BJ093078.1 GI:17592487	
SOURCE	EST.	
ORGANISM	African clawed frog.	
	Xenopus laevis	
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;	
	Xenopodidae; Xenopus.	
	1 (bases 1 to 646)	
REFERENCE	Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara	
AUTHORS	Y.	
TITLE	Expressed genes in X. laevis embryo	
JOURNAL	unpublished (2001)	
COMMENT	Contact: Tadasi Shin-I	
	Center For Genetic Resource Information	
	National Institute of Genetics	
	1111 Yata, Mishima, Shizuoka 411-8540, Japan	
	Tel: 81-559-81-6856	
	Fax: 81-559-81-6855	
	Email: tshin@genes.nig.ac.jp.	
FEATURES	Location/Qualifiers	
source	1..646	
	/organism="Xenopus laevis"	
	/db_xref="taxon:8355"	
	/clone="X141118"	
	/clone_1lb="N1B Mochii normalized Xenopus early gastrula	
	library"	
	/tisue-type="whole embryo"	
	/dev-stage="stage 10.5"	
	/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; CDNAS	
	were oligo-dr primed and directionally cloned. Staging	
	according to MueenKoop and Faber. Library is substracted	
	and was constructed by N. Garrett and A.M. Zorn,	
	(Wellcome/CRC Institute)."	
BASE COUNT	208 a 129 c 161 g 148 t	
ORIGIN		

<b>Alignment Scores:</b>					
Pred. No.:	3,16e-91	Length:	646		
Score:	831.50	Matches:	169		
Percent Similarity:	82.93%	Conservative:	1		
Best Local Similarity:	82.44%	Mismatches:	23		
Query Match:	93.85%	Indels:	3		
DB:	13	Gaps:	2		

  

US-09-820-003B-2 (1-173) x BJ093078 (1-646)

OY	1	MetSerSemeIaenProclutrrAsPryLeuPhelysleuleuilegIyAsper	20
Db	35	AATGCGACAGTAAATCCAGTAGTACATTATTTCAATTACTTGATGGAGACTC	94
OY	21	GIYValGIlyserCysleuleuenaRphelaasphThrTyThrGluserTy	40
Db	95	GGGTGGGAAGTGTCTCTCTGCTTGGATTGGANGVGCACATACAGAAACTTC	154
OY	41	IleserThrIlegIyAlaAspPhelysIleatyrThrIleguleuaSpGilystrile	60
Db	155	ATCATGACAATGGTGTGATTATAATCAGACTATAGATGAAAAAATTC	214
OY	61	LysleugInlle-----	64
Db	215	AAACTCCAATTGGACACATCCGCCAGAAAGATTTCGACATTTCATTCAGTTAC	274
OY	65	-----glusePhean	68
Db	275	TACAGGGAGACCACGGGCATCATGTTGTGTATGACGTAAACAGACCAGGAATTCCTCAC	334
OY	69	AsnValLysGIINTPleuGngluIleaspArgTYralaserGIunValasnLyseu	88
Db	335	AATGTCAAGCAGTGGCTTCAGGAATATGATCGTTATGCCAGTGAAGCTTAAAGAATTA	394
OY	89	LeuValGIyAnlYscYkAsplemThrLysLysValAlaSPtyThrThalaly	108
Db	395	TTGGTGGCAACAATGTATCTCACAAAAGAAAGTGTGGATCTACACACAGCAAG	454
OY	109	GIuPhaIlaAspsertleuGIyletrophleuGIurhsertIalaasnaIethaan	128
Db	455	GATTTGGCAGACTCTCTGGGAATTCATTTTTGGAACMACCGCAAGACACACAAAC	514
OY	129	VaIGlnGlnserPheacthrMetAlaIleGluIleLysLysArgmegIyProGlyAla	148
Db	515	GIYGACACAGGCCTTCATGACAAATGGACGGGGAATCAAAGGAATGGCCCTGGGGCC	574
OY	149	ThrAlaGIyAlaIleGIySerasnValylIleGlnserThProvalylsGlnser	168
Db	575	ACAGCAGGGGGTCAAGAGAG--AACGTCAAAATCCAGACACTCATGATCAACATCT	631
OY	169	GIyGIyLycycys 173	
Db	632	AGTGGAGATGCTCC 646	

  

RESULT 7	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	JOURNAL	COMMENT
BQ219698	BQ219698	597 bp mRNA linear EST 02-MAY-2002	BQ219698	BQ219698	IMAGE:6051445 5', mRNA sequence.	house mouse.	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	1 (bases 1 to 957)	NIH-MGC http://mgi.nci.nih.gov/National_Institutes_of_Health_Mammalian_Gene_Collection_(MGC)_Unpublished_(1999)	Contact: Robert Strausberg, Ph.D. Email: cgsabf-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
 Plate: LLM13505 row: e column: 14  
 High quality sequence stop: 671.  
 Location/Qualifiers

## FEATURES

source

1..957  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6051445"  
 /clone\_lib="NCI\_CGAP\_St1"  
 /lab\_host="DH10B (TI-resistant)"  
 /note="Organ: stomach; Vector: pCMV-Sport6; Site: 1: SalI;  
 Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.77 kb. Library constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 269 a 208 c 250 g 229 t 1 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 3,59e-90 Length: 957  
 Score: 825.00 Matches: 171  
 Percent Similarity: 83.41% Conservative: 0  
 Best Local Similarity: 83.41% Mismatches: 2  
 Query Match: 93.12% Gaps: 33  
 DB: 14 Indels: 1

US-09-820-003b-2 (1-173) x BQ129698 (1-957)

OY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
 DB 200 ATGTCACAGCAATCCGAAATATGATATATCAAGTACTCTGATGGCAGTCT 259  
 OY 21 GlyValGlySerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40  
 DB 260 GGGGTGGAAAGTCTGCTCTCTCTAGTTGAGATGATGATACGTAACGAAAGCTAC 319  
 OY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
 DB 320 ATCAGCACAATGGTGTGATTTCAAGATACCACTATAGATTGAGGAAACATC 379  
 OY 61 LysLeuGlnIle----- 64  
 DB 380 AAGCTACAGATATGGGACACAGACGCGAAGATTTCGAACATCATTCCAGTTAT 439  
 OY 65 -----GluSerPheAsn 68  
 DB 440 TACAGAGAGCCCATGCATCATAGTTGTATGATGTGACAGATCAGGAGTCTTCAT 499  
 OY 69 AsnValLysGlnTyrLeuGlnIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88  
 DB 500 AACGTTAAACAGTGGCTCGAGAGATGATCGCTACGCCAGGAAATATCTCAACAAGTTG 559  
 OY 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThaLys 108  
 DB 560 TTGGTAGGGAACAAATGTGACCTGACCAAGAAAGTAGTAGTACACACACAGCAAG 619  
 OY 109 GluPheAlaAspSerLeuGlyTyrLeuPheLeuGluTyrThrAlaLysAsnAlaThrAsn 128  
 DB 620 GAATTTGAGATTCCTTGGAATTCATTGGAACCAAGGCTGAAGACCAACGAT 679  
 OY 129 ValGluGlnSerPheMetThrMetAlaIleGluIleLysArgMetGlyProGlyAla 148  
 DB 680 GTAGAACAGTCTTTCATGACATGCGACCTGAGATTAAACGCAATGGTCTCGAGCT 739  
 OY 149 ThrAlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168  
 DB 740 ACAGCTGTGTGGCCGAGAAATCCAT-GTAAATAACAGAGACATCCAGTCAAGCAATCA 798  
 OY 169 GlyGlyGlyCysCys 173

DB 799 GGTGAGGCTGCTGC 813

RESULT 8  
 BQ437146

LOCUS BQ437146 864 bp mRNA linear EST 24-MAY-2002  
 DEFINITION AGENCOURT\_7909781 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6156281  
 5', mRNA sequence.

ACCESSION BQ437146

VERSION BQ437146.1 GI:21176210

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC <http://mhc.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
 Plate: LLM13500 row: e column: 18  
 High quality sequence stop: 658.  
 Location/Qualifiers

## FEATURES

source

1..864  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6156281"  
 /clone\_lib="NIH\_MGC\_71"  
 /tissue\_type="Telomysarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: uterus; Vector: pCMV-Sport6; Site: 1: NotI;  
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2.1 kb."  
 BASE COUNT 249 a 171 c 231 g 213 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 7.15e-90 Length: 864  
 Score: 822.00 Matches: 170  
 Percent Similarity: 82.93% Conservative: 0  
 Best Local Similarity: 82.93% Mismatches: 2  
 Query Match: 92.78% Indels: 33  
 DB: 14 Gaps: 1

US-09-820-003b-2 (1-173) x BQ437146 (1-864)

OY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
 DB 184 ATGTCACAGCAATCCGAAATATGATATATCAAGTACTCTGATGGCAGTCT 243  
 OY 21 GlyValGlySerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40  
 DB 244 GGGGTGGAAAGTCTGCTCTCTCTAGTTGAGATGATGATACATATACGAAAGCTAC 303  
 OY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
 DB 304 ATCAGCACAATGGTGTGATTTCAAGATACCACTATAGAGTTAGACGGAAACATC 363  
 OY 61 LysLeuGlnIle----- 64  
 DB 364 AAGCTTCAAAATATGGGACACAGACGCGAAGAAATTTCGAACATCACCCTCAGTTAT 423  
 OY 65 -----GluSerPheAsn 68  
 DB 424 TACAGAGAGCCCATGCATCATAGTTGTATGATGTGACAGATCAGGAGTCTTCAT 483



```

OY 69 AasnVallysgIntPrleugIngluileasparGTYralaserGluasnValasnlyseu 88
DB 484 AATGTTAAACAGCGGTGCGAGAAATAGATCGTTATGCGACGAAATGTCACAAATTTG 543
OY 89 LeuValIGlyAsnlyGysAspLeuThrThlyGlyValValaspyrThrThralaYs 108
DB 544 TTGGTAGGACAAACAGTATGATGACACAAACAAAGTAGTACACACACAGAGAG 603
OY 109 GluPhealaasPserleuGlyleProPheleuGluThrSerAlaYasanaIathrasn 128
DB 604 GAATTCGTATTCCTTGGAATTCCTTTGGAAACCAAGCTTAAGAATGCAACGAT 663
OY 129 ValGluInserPheMetThrMetAlaAlaGluIleYlyAsArgMetGlyProGlyAla 148
DB 664 GTACAGACAGCTTTTCATGACGATGCGACGACGACATTTAAACCAATGGCTCCGACACA 723
OY 149 ThrAlaGlyAlaIleGlySerAsnVallyleGInserThrProVallysgInser 168
DB 724 ACAGCTGGTGTGCTGAGACGATGCAATGTTAAATTCAAGACACTCCAGTCAAGCAGTCA 783
OY 169 Gly-GlyGlyCys 172
DB 784 GGTGGAGAGTTGC 796

RESULT 9
BO944368 901 bp mRNA linear EST 21-AUG-2002
LOCUS BO944368
DEFINITION IMAGE:6441874.5', mRNA sequence.
ACCESSION BO944368
VERSION BO944368.1 GI:22359846
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 901)
NIR-MGC http://mgc.nci.nih.gov/.
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AMJ3965 row: m column: 11
High quality sequence stop: 662.

FEATURES
source location/Qualifiers
1 901
/organism="Mus musculus"
/strain="FVB/N-3"
/old_xref="taxon:10090"
/clone="IMAGE:6441874"
/clone_1lb="NCL_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by: Life Technologies. Investigator
providing samples: Gilbert Smith, NIR"

BASE COUNT 246 a 197 c 243 g 214 t 1 others
ORIGIN

```

```

Best Local Similarity: 82.13% Mismatches: 2
Query Match: 92.55% Indels: 34
DB: 14 Gaps: 1
US-09-820-003b-2 (1-173) x BO944368 (1-901)

OY 1 MetSerMetLaspProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
DB 188 ATGTCAGCATGAATCCCGAATATGATATTTATTCAGTTACTTGTGATGGCATTC 247
OY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaaspAspThrThrGluSerTyr 40
DB 248 GGGGTGGAAATCTCTGCTTCCTTCTAGGTTGCAATATACGATACGGAAGATC 307
OY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuaspGlyLysThrIle 60
DB 308 ATCAGCACAAATGTGTGATTCAGATACGACACTATACGTTACGATGGGAAACATC 367
OY 61 LysLeuGluIle----- 64
DB 368 AAGCTACAGATATGGGACACAGACAGGCCAGAAAGATTGCAACATCATTCCACTAT 427
OY 65 -----GluSerPheasn 68
DB 428 TACAGAGAGCCCATGCGCATCATGTTGTTGATGTGACACATGAGAGCTTCAT 487
OY 69 AasnVallysgIntPrleugIngluileasparGTYralaserGluasnValasnlyseu 88
DB 488 AACTTTAAACAGTGTGCTGACAGAGATGATCGCTACGCCAGTAAATGTTCACAAAGTTG 547
OY 89 LeuValIGlyAsnlyGysAspLeuThrThrThlyGlyValValaspyrThrThralaYs 108
DB 548 TTGGTAGGACAAACAGTATGATGACCTGACCAAGAAAGTAGTACACACACAGCAAG 607
OY 109 GluPhealaasPserleuGlyleProPheleuGluThrSerAlaYasanaIathrasn 128
DB 608 GAATTCGAGATTCCTTGGAATTCATTTTGGAAACCAAGTCTCAGACGCAACCAAT 667
OY 129 ValGluInserPheMetThrMetAlaAlaGluIleYlyAsArgMetGlyProGlyAla 148
DB 668 GTACAGACAGCTTTTCATGACGATGCGACGATGATTTAAACGATGGCTCGAGAGCT 727
OY 149 ThrAlaGlyAlaIleGlySerAsnVallyleGInserThrProVal-LysAlaIas 168
DB 728 ACAGCTGGTGTGCGCGAGAGTCCATGTTAAATCCAGAGNACCTCCAGTCAAGCAGC 787
OY 168 ergGlyGlyGlyCysCys 173
DB 788 CAGGTGAGGCTGCTGC 804

RESULT 10
BM452262 1105 bp mRNA linear EST 05-FEB-2002
LOCUS BM452262
DEFINITION AGENCOURT_6386191 NIR_MGC_72 Homo sapiens cDNA clone IMAGE:5526535
5', mRNA sequence.
ACCESSION BM452262
VERSION BM452262.1 GI:18501302
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1105)
NIR-MGC http://mgc.nci.nih.gov/.
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCFD/DTF
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

```

Alignment Scores: 1,34e-89 Length: 901  
Pred. No.: 820.00 Matches: 170  
Score: 82.618 Conservative: 1  
Percent Similarity: 82.618

found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM12200 row: j column: 12  
 High quality sequence stop: 733.  
 Location/Qualifiers  
 1..1105

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Nih\_MGC\_72"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;  
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2 kb. Library constructed by Life  
 Technologies."

BASE COUNT 320 a 216 c 279 g 282 t 8 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,43e-89 Length: 1105  
 Score: 819.00 Matches: 172  
 Percent Similarity: 83.50% Conservative: 0  
 Best Local Similarity: 83.50% Mismatches: 1  
 Query Match: 92.44% Indels: 34  
 DB: 13 Gaps: 1

US-09-820-003b-2 (1-173) x BM452262 (1-1105)

OY 1 MetSerSerMetAspProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
 DB 182 ATGTCACGATGATATCCGAAATATGATTTATTCACATGTTCTGTGATGGCGACTCA 241  
 OY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40  
 DB 242 GGGGTGGAAAGCTCTGCTCTCTTGTAGTTGGAGATGATACATATACAGAAAGCTAC 301  
 OY 41 IleserThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
 DB 302 ATCAGACACAAATGGTGTGATTTCAAAATAGACATATAGATTAGCGGAAACAAATC 361  
 OY 61 LysLeuGlnIle----- 64  
 DB 362 AAGCTTCAAATATGGGACACAGCAGGCCAGAAAGATTTCGAACATCACTCCAGTTAT 421  
 OY 65 -----GluSerPheasn 68  
 DB 422 TACAGAGAGAGCCCATGGCATCATAGTTGTATGATGTGACAGATCAGAGAGTCCCTCAAT 481  
 OY 69 AsnValLysGlnTyrPheGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88  
 DB 482 AATGTTAAACACTGGCTGAGGAAATAGATCCTTATGCCAGTGAATAATGTCAACAAATG 541  
 OY 89 LeuValGlyAsnLysCysAspLeuThrThrLysValValAspTyrThrThrAlaLys 108  
 DB 542 TTGTAAGGACAAATATGATCTGACACCAAGAAAGTAGTAGCTACCAACAGGAAAG 601  
 OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128  
 DB 602 GAATTTGGTGAATCCCTTGGAAATCCGTTTGGAAACAGTGCATAGAGATGCACAGAT 661  
 OY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148  
 DB 662 GTAGAACAGTCTTTCATGACATGACACTGAGATTAAAGCAATGGGTCGCCGAGACA 721  
 OY 149 ThrAlaGlyValAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168  
 DB 722 ACAGCTGGGTGGTGGAGAA-TCCAATGTTAAATTCAGAGACATCCAGTCAAGACATCA 780  
 OY 169 -GlyGlyLysCysCys 173  
 DB 781 NGGTGAGGTTCCTGCTGC 796

RESULT 11  
 AM954378 690 bp mRNA linear EST 01-JUN-2000  
 LOCUS EST166448 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.  
 DEFINITION  
 AM954378  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 690)  
 Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt  
 I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
 Quackenbush,J.  
 Assessment of gene expression patterns in a model of colon tumor  
 metastasis using a 19,200 element cDNA microarray  
 Unpublished (2000)  
 CONTACT  
 JOURNAL  
 COMMENT  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 3528  
 Fax: 301 838 0208  
 Email: john@etlgr.org  
 Plate: 63  
 Seq primer: Reverse.  
 Location/Qualifiers  
 1..690

## FEATURES

source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="MAGE resequences, MAGE"  
 /note="Vector: pBluescriptSK"

BASE COUNT 219 a 130 c 168 g 173 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 4.81e-89 Length: 690  
 Score: 814.00 Matches: 169  
 Percent Similarity: 83.25% Conservative: 0  
 Best Local Similarity: 83.25% Mismatches: 1  
 Query Match: 91.87% Indels: 33  
 DB: 10 Gaps: 1

US-09-820-003b-2 (1-173) x AM954378 (1-690)

OY 1 MetSerSerMetAspProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
 DB 31 ATGTCACGATGATATCCGAAATATGATTTATTCACATGTTCTGTGATGGCGACTCA 90  
 OY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40  
 DB 91 GGGGTGGAAAGCTCTGCTCTCTTGTAGTTGGAGATGATACATATACAGAAAGCTAC 150  
 OY 41 IleserThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
 DB 151 ATCAGACACAAATGGTGTGATTTCAAAATAGACATATAGATTAGCGGAAACAAATC 210  
 OY 61 LysLeuGlnIle----- 64  
 DB 211 AAGCTTCAAATATGGGACACAGCAGGCCAGAAAGATTTCGAACATCACTCCAGTTAT 270  
 OY 65 -----GluSerPheasn 68  
 DB 271 TACAGAGAGAGCCCATGGCATCATAGTTGTATGATGTGACAGATCAGAGAGTCCCTCAAT 330  
 OY 69 AsnValLysGlnTyrPheGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88  
 DB 331 AATGTTAAACACTGGCTGAGGAAATAGATCCTTATGCCAGTGAATAATGTCAACAAATG 390  
 OY 89 LeuValGlyAsnLysCysAspLeuThrThrLysValValAspTyrThrThrAlaLys 108  
 DB 391 TTGTAAGGACAAATATGATCTGACACCAAGAAAGTAGTAGCTACCAACAGCAGAG 450

109 GluphealaspserleuglylleprophaleugluThrsSerAlaLysAsnAlaThrAsn 128  
 DB |||||||  
 451 GAATTCCTATTCCTTCGAAATTCCTTTTGCACCACTGCTAGATGCAACGAAAT 510  
 QY |||||||  
 129 ValGluInserPheMetThrMetAlaAlaGluIleLysArgMetGlyProGlyAla 148  
 DB |||||||  
 511 GTAAGACAGCTTCCTCAATGAGATGAGCGAGATTAATAAGCAATGAGTCCGAGACA 570  
 QY |||||||  
 149 ThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThr-ProValLysGlnSe 168  
 DB |||||||  
 571 ACACGGGGTGGTCCCGAATAATCCAAATGTTAAATTCAGACACTTCAGTCAACAGCTC 630  
 QY |||||||  
 168 IGLyly 170  
 DB 631 AGGTGG 637  
 RESULT 12  
 AL530265 945 bp mRNA linear EST 13-FEB-2001  
 LOCUS AL530265 LTI\_NFL001\_NBC4 Homo sapiens cDNA clone CSDD009YH19 5  
 DEFINITION Prime, mRNA sequence.  
 ACCESSION AL530265  
 VERSION AL530265.1 GI:12793758  
 KEYWORDS EST  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 945)  
 AUTHORS Li, W.B., Gruber, C., Jessup, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seque@genoscope.cns.fr, web : www.genoscope.cns.fr.  
 FEATURES  
 source  
 1. 945  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CSDD009YH19"  
 /clone\_lib="LTI\_NFL001\_NBC4"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer, five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed  
 by Life Technologies. Contact : Peng Huang Life  
 Technologies, a division of Invitrogen 9800 Medical Center  
 Drive Rockville, Maryland 20850, USA fax : (1) 301 610  
 8371 Email : liliang@litech.com URL :  
 http://fulllength.invitrogen.com"  
 BASE COUNT 280 a 186 c 224 g 246 t 9 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.19e-88 Length: 945  
 Score: 809.00 Matches: 171  
 Percent Similarity: 83.01% Conservative: 0  
 Best Local Similarity: 83.01% Mismatches: 2  
 Query Match: 91.31% Indels: 34  
 DB: 9 Gaps: 1  
 US-09-820-003b-2 (1-173) x AL530265 (1-945)  
 QY 1 MetSerGlnMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
 DB |||||||  
 121 ATGTCACGACGATGAATTCGCAATATGATATATTCATCAAGTTACTTCGATCGCACACA 180  
 QY 21 GlyValGlySerCysLeuLeuLeuAlaGlnPheAlaAspThrTyrThrGlnSerTyr 40

DB |||||||  
 181 GGGGTGGAAATCTCTGCTCTTCTTTCAGATGATACATACAGAACTCTAC 240  
 QY |||||||  
 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
 DB |||||||  
 241 ATCAGCAATATGGTGTGGATTTCAAAATAGAACTATAGATAGACGGGAAACAAATC 300  
 QY |||||||  
 61 LysLeuGlnIle----- 64  
 DB |||||||  
 301 AAGCTCAAAATATGGACACAGCAGCGCCAGCAAAAGATTTCGAACATCCCTCAGTAT 360  
 QY |||||||  
 65 ----- GluSerPheAsn 68  
 DB |||||||  
 361 TACAGAGAGCCCATCGCATCATAGTGTGTATGATGTGACAGATCAGAGACTCTTCAT 420  
 QY |||||||  
 69 AsnValLysGlnTyrPleuGlnGluIleAspArgTyrIleSerGluAsnValAsnLysLeu 88  
 DB |||||||  
 421 AATGTTAAACAGTGGCTGCGAGAAATAGATGTTATCCAGAGAAATGTCACAAATG 480  
 QY |||||||  
 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108  
 DB |||||||  
 481 TTGGTAGGGAACAAATGTGATCTGACCAAAAGAAAGTAACTACACAAACAGCAAG 540  
 QY |||||||  
 109 GluphealaspserleuglylleprophaleugluThrsSerAlaLysAsnAlaThrAsn 128  
 DB |||||||  
 541 GAATTCGCTGATTCCTTCGAAATTCGTTTGGAAACAGCTGTAAGAAAGCAACAGCA 600  
 QY |||||||  
 129 ValGluInserPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148  
 DB |||||||  
 601 GTGAACACTCTTCATGACGATGAGCGAGCTGAGATTAATAAGCAATGGGTCCGGA-GCA 659  
 QY |||||||  
 149 ThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThr-ProValLysGlnSer 168  
 DB |||||||  
 660 ACAGCTGGGGTGGTCCCGAATAATCCAAATGTTAAATTCAGACACTTCAGTCAACAGCTCA 719  
 QY |||||||  
 169 -GLYGLYGLYCysCys 173  
 DB |||||||  
 720 GGGTGGAGGCTGCTGC 735  
 RESULT 13  
 B073109 642 bp mRNA linear EST 11-DEC-2001  
 LOCUS B073109 NIBB Mochii normalized Xenopus tailbud library Xenopus  
 DEFINITION laevis cDNA clone XL108p01 5', mRNA sequence.  
 ACCESSION B073109  
 VERSION B073109.1 GI:17503238  
 KEYWORDS EST  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
 Xenopodidae; Xenopus.  
 REFERENCE 1 (bases 1 to 642)  
 AUTHORS Kiteyama, A., Terasaka, C., Mochii, M., Deno, N., Shin-I, T. and Kohara,  
 Y.  
 TITLE Expressed genes in X. laevis embryo  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Tadasu Shin-I  
 Center for Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp.  
 FEATURES  
 source  
 1. 642  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone="XL108p01"  
 /clone\_lib="NIBB Mochii normalized Xenopus tailbud  
 library"  
 /tissue\_type="whole embryo"  
 /dev\_stage="stage 25"



Db 626 GGAATTGCAGATTCCTTGGATTCATTCATTTGGAAACAGCTGCTAACAAGCAACAA 685  
OY 128 nvaIgluIserPheMetThrMetAlaIaIaIuIleLysArgMet-GlyProGlyA 148  
Db 686 TGTGAAACACTCTTTATGACGATGCGAGCTGATTAAGGAAATGGGCTCCTGAG 745  
OY 148 IatPrIaIaIyGlyIaIaIuLys-SerAsnValLysIleGlnSerThrProValGln 167  
Db 746 CTACAGCTGCTGCTGCCGAGAAAGCTCCAAATGTTAAATCCAGACCTCCATCAGACG 805  
OY 168 SerGlyIyGlyCys 172  
Db 806 TCAGGTGAGGCTGC 820  
RESULT 15  
LOCUS B1739625 855 bp mRNA linear EST 20-SEP-2001  
DEFINITION 603361882F1 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:536876 5',  
mRNA sequence.  
ACCESSION B1739625 GI:15716638  
VERSION B1739625  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 855)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Ph.D.  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: rgs@fremail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIML at:  
http://image.llnl.gov  
Place: LHAM11938 row: d column: 21  
High quality sequence atop: 849.  
Location/Qualifiers  
1. 855  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:536876"  
/clone\_lib="NIH\_MGC\_94"  
/issue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1; Note:  
Site 2: Salt; Cloned unidirectionally; oligo-dT primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC library."  
BASE COUNT 234 a 181 c 233 g 205 t 2 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,36e-88 length: 855  
Score: 806.00 Matches: 169  
Percent Similarity: 82.218 Conservative: 2  
Best Local Similarity: 81.258 Mismatches: 35  
Query Match: 90.978 Indels: 13  
Gaps: 1  
US-09-820-003b-2 (1-173) x B1739625 (1-855)  
OY 1 MetSerSerMetLeuAnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20  
Db 185 ATGTCCAGCAGATGAAATCCCAATATGATTAATTTATTCAGTTACTTCGATGGCGATTCT 244  
OY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGlnSerTyr 40

Db 245 GGGATTGAAATCCTCGCTTCTCCTTAGCTTCAGATGATACGATATACGAAACCTAC 304  
OY 41 IIsSerThrIleGlyValAspPheLysIleArgThrIleGlnLeuAsnGlyLysThrIle 60  
Db 305 ATCGACAACTTGGGTGATGATTTCAAGATACGAACTATAGTAGATGGAAACAAATC 364  
OY 61 LysLeuGlnIle----- 64  
Db 365 AAGCTACATATGTGGACACAGACGACGAGAAAGATTTTGCAACATACCTCCAGTTAT 424  
OY 65 -----GluSerPheAsn 68  
Db 425 TACAGAGAGACCCATGCATCATAGTTGTATGATGTGACAGATCAGAGCTCTTCATAT 484  
OY 69 AsnValLysGlnTrpLeuGlnIuIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88  
Db 485 AAGCTTAAACAGTGGCTGCGAGAGATGATGCTACGCCAGTGAATAATGTCACACAGTNN 544  
OY 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrAlaLys 108  
Db 545 GTGGTAGGGAACAATGTGACCTGACCAACAAAGAAAGTACTAGACTACACACACAGAAAG 604  
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128  
Db 605 GAAATTGCAGATTCCTTGGATTCATTTTGGAAACAGCTCTAAGACGACGAAGAT 664  
OY 129 ValIgluIserPheMetThrMetAlaIaIaIuIleLysArgMetGly--ProGlyA 148  
Db 665 GTGAAACACTCTTTATGACGATGCGAGCTGATTAAGGAAATGGGCTCCTGAG 724  
OY 148 IatPrIaIaIyGlyIaIaIuLysSerAsnValLysIleGlnSerThrProValGln 168  
Db 725 CTACAGCTGCTGCTGCCGAGAAAGCTCCAAATGTTAAATCCAGACCTCCATCAGACG 784  
OY 168 SerGly-GlyGlyCysCys 173  
Db 785 CAGGTGAGGCTGCTGC 802

Search completed: June 22, 2003, 18:55:41  
Job time: 1427 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 22, 2003, 16:26:39 ; Search time 216 Seconds  
(without alignments)  
1803.684 Million cell updates/sec

Title: US-09-820-003b-2

Perfect score: 86  
Sequence: 1 MSSKNPEYDYLFKLLIGDS.....EKSNNKIQSTPVKSGGGCC 173

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Command line parameters:  
MODE=frame\_plus.model -DEV=klp  
-O=/cgn2\_1/us0982003/runat\_18062003\_144129\_13082/app\_query.fasta\_1.3227  
-DB=N\_Geneseq\_101002 -QPM=fastap -SOFFIT=p2n.rng -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -ONITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdl  
-LIST=45 -DOCALLGM=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0982003.ecgn.1.1.396.etunat.18062003.144129.13082 -NCPU=6 -ICPU=3  
-NO\_MMAP -LANG=ENGLISH -NEG\_SCORES=0 -WAIT -DSPBLLOC=100 -LONGLOC  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :  
N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	860	97.1	723	24	ABK83978		Human cDNA differe
2	860	97.1	2528	21	AC60009		Human secreted pro
3	768.5	86.7	925	20	AAZ2732		Human Rab protein,
4	768.5	86.7	939	21	AAZ2161		Human breast and o
5	768.5	86.7	1898	22	AA544924		cDNA encoding nove
6	768.5	86.7	1944	22	AA193456		Human polynucleoti
7	733.5	82.8	1965	22	AA534668		Human DNA for a no
8	683	77.1	1791	23	ABLO5983		Drosophila melanog
9	657.5	74.2	1202	21	AAFL14053		Aspergillus oryzae
10	639	72.1	959	21	AAAC33987		Arabidopsis thalia
11	633	71.4	932	21	AAAC35200		Arabidopsis thalia
12	624	70.4	1193	23	AAAC38600		Arabidopsis thalia
13	619.5	69.9	777	21	AAAC42684		Arabidopsis thalia
14	619.5	69.9	881	24	ABN98761		Arabidopsis thalia
15	612	69.1	666	21	AAAC42764		Arabidopsis thalia
16	612	69.1	932	21	AAAC48274		Arabidopsis thalia
17	597	67.4	607	22	AAH87926		Arabidopsis thalia
18	580	65.5	594	22	AAAC49581		peppermint plant o
19	570	64.3	1380	21	ABO5982		Arabidopsis thalia
20	529	59.7	5019	23	ABO5982		Fusarium venenatum
21	460	51.9	749	21	AAFL14316		Drosophila melanog
22	452	51.0	615	22	AAH87928		Trichoderma reesei
23	444	50.1	911	21	AAAC38429		peppermint plant o
24	442	49.9	1101	21	AAAC34080		Arabidopsis thalia
25	440.5	49.7	1203	21	AAAC34080		Arabidopsis thalia
26	431	48.6	1129	21	AAAC44482		Arabidopsis thalia
27	429.5	48.5	705	21	AAAC51491		Arabidopsis thalia
28	428.5	48.4	1023	21	AAAC54347		Human rab8 homolog
29	428.5	48.4	1025	21	AAAC54347		Arabidopsis thalia
30	419	47.3	1866	22	ABO59160		Human ovarian anti
31	417	47.1	2411	22	AAH75182		Nucleotide sequenc
32	414	46.7	1161	22	AAH75182		Human cancer agent
33	414	46.7	2497	22	AA560884		Human cancer agent
34	414	46.7	2497	22	AA560884		Human cancer agent
35	414	46.7	2497	22	AA560884		Human cancer agent
36	414	46.7	2497	22	AA560884		Human cancer agent
37	414	46.7	2497	22	AA560884		Human prostate exp
38	414	46.7	2497	22	AA560884		Human prostate exp
39	414	46.7	2497	22	AA560884		Human prostate exp
40	414	46.7	2497	22	AA560884		Human prostate exp
41	413	46.6	866	22	AAH13912		Human cDNA sequenc
42	408	45.6	2247	23	ABO29651		Human cDNA clone (
43	405	45.7	1492	21	AAAC38136		Zea mays DNA fragm
44	403.5	45.5	559	21	AAAC44411		Arabidopsis thalia
45	403.5	45.5	716	21	AAAC40104		Human Radio CDNA
			861	21	AAA40108		Human Radio CDNA

## ALIGNMENTS

RESULT 1  
ABK83978  
ID ABK83978 standard; cDNA: 723 BP.  
XX  
AC ABK83978;  
XX  
DT 14-AUG-2002 (first entry)  
XX  
DE Human cDNA differentially expressed in granulocytic cells #549.  
XX  
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
KW viral infection; parasitic infection; protozoal infection;  
KW fungal infection; sterile inflammatory disease; psoriasis;  
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
KW adult respiratory distress syndrome; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; peritoneal disease;  
KW granulocyte activation; chronic inflammation; allergy.

XX	
OS	Homo sapiens.
XX	
PN	MO200228999-A2.
XX	
PD	11-APR-2002.
XX	
PF	03-OCT-2001; 2001WO-US30821.
XX	
PR	03-OCT-2000; 2000US-237189P.
XX	
PA	(GENE-) GENE LOGIC INC.
XX	
PI	Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX	
DR	WPI; 2002-435328/46.
XX	
PT	Detecting granulocyte activation by detecting differential expression
PT	of genes associated with granulocyte activation, which serves as
PT	diagnostic markers that is useful for monitoring disease states and
PT	drug toxicity
XX	
ES	Claim 1; SEQ ID No 549; 11app; English.

The invention relates to detecting (M1) granulocyte (GC) activation by (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g., psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

Alignment Scores:		
Pred. No.:	4,33e-89	Length: 723
Score:	860.00	Matches: 173
Percent Similarity:	84.39%	Conservative: 0
Best local Similarity:	84.39%	Mismatches: 0
Query Match:	97.07%	Indels: 32
DB:	24	Gaps: 1
US-09-620-0038-2 (1-173)	XBR63978 (1-723)	

OY		1	MeserSerMetIaPnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer	26
Db		51	ATGCCAGCATGAATCCCGAATATGATTATTATTCACGTACTCTTGATTGGCAGCTCA	11
OY		21	GlyValGlyLysSerCysLeuLeuLeuArgPheLeuLaaspasphrThrTrhgInserTyr	40
Db		111	GGGGTTGGAAAGCTTGCTGCTCTCTTAAGGTTGCACATGATACATATACAGAAGTAC	11
OY		41	IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle	60
Db		171	ATCGCACAAATTGGGTGTGATTTCCAATAATAGAACTATAGAGTTAGCGGGAAAACATC	22
OY		61	LysLeuGlnIle-----	6-
Db		231	AAGCTCAAAATATGGGACACAGCAGGCCAGGAAAGATTGGAACATCACTCCACGTTAT	28
OY		65	-----GluSerPheAsn	66
Db		291	TACAGAGAGGCCATGGCATCATGATGTTGTGTATGATGTGACAGATCGAGAGTCTTCAT	35
OY		69	AsnValLysGlnTrpLeuGlnGluIleAsparGlyAlaSerGluAsnValAsnLysLeu	86
Db		351	AATGTTAAACAGTGGCTGCAGGAATAGATCGTTATCCAGTGAATGTCAAACAATWG	41
OY		89	LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspyrThrThrAlaLys	110
Db		411	TTGGTAGGGAACCAATGTGATCTGACCACAAAGAAATAGATAGACTACACACAGCCAG	47
OY		109	GluPheLeuLaaspSerLeuGlyIleProPheLeuGlnThrSerAlaLysAsnAlaThrAsn	117
Db		471	GAATTTGCTGATTCCTCTTGGAAATTCGGTTTTTGGAAACAGTGGTAAATGCAAGAAC	53
OY		129	ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla	140
Db		531	GTAACACAGTCTTTCATGACAGATGGCAGCTGAGATTAAAAAGCAATGGGTCCCGGAGCA	59
OY		149	ThrIlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer	160
Db		591	ACACCTGTGTGGTGTGGAAGTCCAAATGTTAAATTTCAGAGCACTCCAGTCAAGCACTCA	65
OY		169	GlyGlyLysCysCys	173
Db		651	GGTGGAGTGTGCTGC	665
<b>RESULT 2</b>				
AAC60009	ID	AAC60009	standard; cDNA; 2528 BP.	
XX	XX	AAC60009;		
XX	XX	26-JAN-2001	(first entry)	
DE	XX	Human secreted protein gene 44	SEQ ID NO:54.	
KW	XX	Human; secreted protein; diagnosis; cytostatic; immunosuppressive;		
KW	XX	nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;		
KW	XX	antidiabetic; antiinflammatory; anticancer; vulnerrary; anticonvulsant;		
KW	XX	antibacterial; antifungal; antiparasitic; cardiant; gene therapy;		
KW	XX	cancer; immune disorder; cardiovascular disorder; wound healing;		
KW	XX	neurological disease; infectious disease; chromosome identification; ss		
OS	XX	Homo sapLens.		
PN	XX	WO200058356-A1.		
PD	XX	05-OCT-2000.		
PF	XX	22-MAR-2000; 2000WO-USO7535.		
PR	XX	26-MAR-1999; 99US-0126511.		
PR	XX	17-DEC-1999; 99US-0172413.		
RA	XX	(HDMA-) HUMAN GENOME SCI INC.		

(HUMA-) HUMAN GENOME SCI INC.





CC treat a disorder associated with cell proliferation e.g. cancers or  
 CC inflammation, e.g. Addison's disease, adult respiratory distress  
 CC syndrome, allergies, asthma, atherosclerosis, bronchitis, cholecystitis,  
 CC Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis,  
 CC diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis,  
 CC gout, Graves' disease, hypereosinophilia, irritable bowel syndrome, lupus  
 CC erythematous, multiple sclerosis, myasthenia gravis, inflammation,  
 CC osteoarthritis, osteoporosis, pancreatitis, polyomyelitis, rheumatoid  
 CC arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis,  
 CC complications of cancer, haemodialysis, extracorporeal circulation,  
 CC infections and trauma.  
 XX

SO Sequence 925 BP; 198 A; 263 C; 288 G; 176 T; 0 other;

#### Alignment Scores:

Pred. No.:	1.78e-78	Length:	925
Score:	768.50	Matches:	156
Percent Similarity:	80.39%	Conservative:	8
Best Local Similarity:	76.47%	Mismatches:	7
Query Match:	86.74%	Indels:	33
DB:	20	Gaps:	2

US-09-820-003B-2 (1-173) x AAF21661 (1-925)

```

OY 2 SerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGly 21
DB 60 GCCGCCATGAACCCGGAATATGACTACCTGTTAAAGCTCTTTGATGGCGACTCAGGC 119
OY 22 ValGlyLysSerCysLeuLeuLeuArpPheAlaAspThrTyrThrGluSerTyrIle 41
DB 120 GTGGGCATGATGCTGCTCCTCGGTTGCTGATGACACTACACAGAGACATACATC 179
OY 42 SerThrIleGlyValAlaAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys 61
DB 180 AGCACCAATCGGGGTGGACTTCATGAATCCGAAACATCGAGCTGGAGAAATCATATCAA 239
OY 62 LeuGluIle----- 64
DB 240 CTTGAGATCTGGGACACAGCGGGCCAGAACGGCTCCGACCATCTCCAGTACTATAC 299
OY 65 -----GluSerPheAsnAsn 69
DB 300 CGGGGGCTCATGCGCATCGTGTGATGACGTCACTGACACAGAAATCTTACCCCAAC 359
OY 70 ValLysGluThrPheGluGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeu 89
DB 360 GTGAAGCAGTGGCTGACAGAGATTGACCGCTATGCCAGCGAGAACGTCAATAGCTCTG 419
OY 90 ValGlyAsnLysCysAspLeuThrThrLysLysValAlaAspTyrThrThrAlaLysGlu 109
DB 420 GTGGGCACCAAGAGCGACCTCACCCAGAGAGGTGGTGGACACACACACAGCCAAAGAG 479
OY 110 PheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnVal 129
DB 480 TTTGGAGACTCTCTGGGACATCCCTCTTGGAGAGACGGCCAGAAATAGCCCAATATGC 539
OY 130 GluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThr 149
DB 540 GAGCAGGGCTTCAATGACCATGCTGCTGAAATCAAAAAGCGATGGGCGCTGGACACCC 599
OY 150 AlaGlyGlyAlaGluThrSerAsnValLysIleGlnSerThrProValLysGlnSerGly 169
DB 600 TCTGGGGGCG---GAGCGGCCCAATCTCAAGATCGACACACCCCTGTAAAGCGGCTGCG 656
OY 170 GlyGlyCysCys 173
DB 657 GGTGGCTCTTGC 668

```

#### RESULT 4

AAF21661  
 ID AAF21661 standard; DNA; 939 BP.  
 XX  
 AC AAF21661;

XX 27-MAR-2001 (first entry)  
 XX Human breast and ovarian cancer associated antigen gene SEQ ID 48.  
 DE  
 XX Human; breast cancer; ovarian cancer; cytoskeletal; immunosuppressive;  
 KW neutrotropic; neuroprotective; antiviral; anti-allergic; hepatotropic;  
 KW antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;  
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;  
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
 KW cardiovascular disorder; wound healing; neurological disease; ds.  
 KW  
 XX Homo sapiens.  
 OS  
 XX MO200055173-AL.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 08-MAR-2000; 2000WO-US05881.  
 PF  
 XX 12-MAR-1999; 99US-0124270.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Rosen CA, Ruben SM;  
 PI  
 XX WPI: 2000-611515/58.  
 DR P-PSDB; AAB58758.  
 XX  
 XX New human breast and ovarian cancer associated gene sequences and the  
 PT polypeptides encoded by these genes, useful in the prevention,  
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular  
 PT disorders and neurological diseases -  
 PT  
 XX  
 XX Claim 1; Page 515; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
 CC associated with breast and ovarian cancer. Included in the invention are  
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
 CC isolation and characterisation of the DNA and protein sequences of the  
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;  
 CC neutrotropic; neuroprotective; antiviral; anti-allergic; hepatotropic;  
 CC antidiabetic; antiparasitic; antitumor; vulnery; anticonvulsant;  
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The  
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,  
 CC particularly breast and ovarian cancer. The nucleic acid sequences,  
 CC proteins, agonists and antagonists may also be used in the diagnosis,  
 CC prevention and treatment of immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; cardiovascular disorders such as  
 CC myocardial ischaemia; wound healing; neurological diseases such as  
 CC cerebral anoxia and epilepsy; and infectious diseases.  
 CC  
 XX

SO Sequence 939 BP; 202 A; 264 C; 289 G; 180 T; 4 other;

#### Alignment Scores:

Pred. No.:	1.82e-78	Length:	939
Score:	768.50	Matches:	156
Percent Similarity:	80.39%	Conservative:	8
Best Local Similarity:	76.47%	Mismatches:	7
Query Match:	86.74%	Indels:	33
DB:	21	Gaps:	2

US-09-820-003B-2 (1-173) x AAF21661 (1-939)

```

OY 2 SerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGly 21
DB 64 GCCGCCATGAACCCGGAATATGACTACCTGTTAAAGCTCTTTGATGGCGACTCAGGC 123

```

Oy		22	VAlG1YlSeSeyCSyleauLeuaAqPhellaAspGrThrTyTrHguISerTyrIle	41
Db		124	GTSGGCAAGTCATGTCTGCTCCCGGGTTTCCTCATGACACTCACAGAGACTCATTAC	187
Oy		42	SerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys	61
Db		184	AGCAACATCGGGGTGAATTCAAAGATCGAAKCATGCAGCTGGATGGCAAACTATCAAA	243
Oy		62	LeuGlnIle-----	64
Db		244	CTTAAGATCTGGGACACAGCGGGCCAGGAAGCTTCGGACCATCTCCAGCTACTAC	303
Oy		65	-----GlutSerPheaaNasn	69
Db		304	CGGGGGGCTCATGGCATCATCTGTGTATGACGTCACATGCACAGAAATCTCTACGCCAAC	363
Oy		70	ValIlyGlnTrpLeuGlnGlnIleAspArgTyrAlaSerGluAsnValAsnLysIleLeu	89
Db		364	GTCAGACATGTGGCTCAGAGAGATTGACCGCTATCCAGCGAGAAAGTCATAATAGCTCCG	422
Oy		90	ValGIYAsnLysCYSAAspleuThrThrLysLysValValAspTyrThrThrAlaLysGlu	109
Db		424	GTCGGCCAAACAAGCGACCTCACACCAAGAAGTGCTGTGAAACACCCACAGCCAAGAG	483
Oy		110	PheIlaAspSerIleuGlyIleProPheLeuIunThrSerAlaLysAsnAlaThrAsnVal	123
Db		484	TTCGCAAGCTCTCTGGCAATCCCCTCTTGTGAGACGAGCGCAAAAGTCACCAATATGC	543
Oy		130	GlnGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThr	145
Db		544	GAGGAGGACCTTATGACCATGAGCTGCTGAATTCAAAAGAGAGATGGGCGCTGAGACAGCC	603
Oy		150	AlaGIYAlaGluLysSerAsnValLysIleGlnIserThrProValLysGlnSerGly	165
Db		604	TCGTGGGGGCG---GAGCGGCCCAATCTCAAGATCGACAGCACCCCTGTAAAGCGCGCTGAG	660
Oy		170	GlyGIYAlaCYcys 173	
Db		661	GGTGGCTGTGGC 672	
RESULT 5				
AAS44924	ID	AAS44924 standard; cDNA; 1898 BP.		
XX	AC	AAS44924;		
XX	DT	18-DEC-2001 (first entry)		
DE	CDNA	encoding novel human secretory protein, Seq ID No. 5.		
XX	Human;	secreted protein; arthritis; Crohn's disease; sepsis; shock;		
KM	Ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;			
KM	transgenic animal; Alzheimer's disease; Parkinson's disease; burn;			
KM	amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;			
KM	ulcer; osteoporosis; bone degenerative disorder; periodontal disease;			
KM	gut protection; lung; liver fibrosis; immune deficiency; infection;			
KM	severe combined immunodeficiency; SCID; autoimmune disorder; allergy;			
KM	multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;			
KM	fertility; analgesic; pain; antigen; ss.			
OS	Homo sapiens.			
XX	PN	MO200166689-A2.		
PD	XX	13-SEP-2001.		
XX	PF	05-MAR-2001; 2001MO-USO4942.		
XX	PR	07-MAR-2000; 2000US-0519705.		
PR	19-MAY-2000; 2000US-057454.			
RR	17-JUN-2000; 2000US-0596193.			
RR	14-JUL-2000; 2000US-0616847.			

PR 19-SEP-2000; 2000US-0665363.  
PR 20-OCT-2000; 2000US-0693267.

PX  
XX  
XX (HYSE-) HXSEQ INC.

PA Tang YF., Liu C., Asundi V., Xu C., Mehman T., Ren F., Ma Y., Zhou P.  
PI Zhao QG., Yang-Y., Dermanac RT., Zhang J., Chen R., Xue AJ., Wang U?  
PX WPI: 2001-569934/66.  
DR P-PEDB: AAU2802A.

XX  
XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
PT prepared from various human tissues, for diagnosis and treatment of  
PT cancer, neurological, inflammatory, and autoimmune disorders -  
XX  
XX  
XX

claim 1; SEQ ID NO 5; 10Pp; English.

xx The invention relates to novel isolated human secreted polypeptides (I)  
CC and polynucleotides (II). (I) and (II) are useful for treating  
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
CC involved in increasing haematopoiesis, stem cell survival, bone growth  
CC and remodeling. (I), (II) and modulators of (II) are useful for  
CC prophylaxis or treatment of one or more cancers. (II) is also useful for  
CC creating transgenic animals useful for studying the in vivo activities of  
CC the polypeptide as well as for studying modulators of the polypeptides.  
CC (I) induces the proliferation of neural cells and regeneration of nerve  
CC and brain tissue and is useful for the treatment of central and  
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
CC Parkinson's disease, Huntington's disease, and amputee lateral  
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic  
CC activity, regulation of hematopoiesis and is useful for treating myeloid  
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
CC and for regeneration of bone, cartilage, tendon alignment and/or nerve  
CC tissue growth, and in tissue repair, healing of burns, incisions,  
CC ulcers for treating osteoporosis, osteoarthritis, bone degenerative  
CC disorders, or periodontal disease. Furthermore, (I) is also useful for  
CC gut protection or regeneration and treatment of lung or liver fibrosis,  
CC disorders involving severe combined immunodeficiency (SCID), bacterial or  
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,  
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic  
CC reactions and conditions, such as asthma or other respiratory problems.  
CC In addition, (I) affects biorythms or circadian cycles of rhythms,  
CC fertility, metabolism, catabolism, anabolism, storage or elimination of  
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides  
CC analgesic effects or other pain reducing effects, immunoglobulin like  
CC activity and can act as an antigen in a vaccine composition to raise an  
CC immune response. AAS4420-AAS4595 represent novel human secreted protein  
CC coding sequences of the invention.

xx  
SQ Sequence 1898 BP; 361 A; 608 C; 537 G; 392 T; 0 other;

Alignment Scores:

	Prod. No.: 4.58e-78	Length: 1898
D Score:	768.50	Matches: 156
E Percent Similarity:	80.3%	Conservative: 8
F Best Local Similarity:	76.47%	Mismatch: 33
G Query Match:	86.74%	Indels: 33
H DB:	22	Gaps: 2

US-09-820-003B-2 (1-1/3) x AAS44924 (1-1898)

OY 2 SerSetMetAsnProGluTyrAspTyrDeuPhelysLeuLeuileGlyaspSerGly 21  
DB : :::::  
46 GCGCCATTAACCCCCGGAAATAGACTCCTGTTCATTGTTGAATGGCACCAAGC 105  
OY 22 ValGIyLysSerCIseuLeuLeuArghHeAlAASPAPfTYrThrGlusertylle 41  
Db GTGCAGCATATATCCCGCCTCCGCTTGCTTCGATGACAGTACAGACAGACTAACTC 165  
OY 42 SetThrlleGlyValAlapheLySllekrThrlleGluleAuSpGlyLysThrlleys 61

```

Db 166 AGCACCATCGGGGTGACTTCAAGATCCGACCATCGACTGGATGGCAAAATATCAAA 225
Oy 62 LeuGlnIle----- 64
Db 226 CTTGAGATCTGGGACACAGCGGGCCAGAGAGGTTCCGAGCATCTCCAGCTACTAC 285
Oy 65 -----GluserPheasnAsn 69
Db 286 CGGGGGGCTCATGGCATCTGCTGTATGACGTACTGACCAAGAAATCTTACCCCAAC 345
Oy 70 VallysgIntPLeuGlnIleAspArgTyrAlaSerGluAsnValAsnIysLeu 89
Db 346 GTGAAGCAGTGGCTCAGAGATTACCGCTATGCCAGCAGAACTCAATAGCTCCTG 405
Oy 90 ValGlyAsnIysCysAspLeuThrThrIysIysValValAspTyrThrAlaIysGlu 109
Db 406 GTGGGCAACAGAGAGGAGCTCCACCAAGAAAGTGTGGACACACACACAGAGAG 465
Oy 110 PheAlaAspSerLeuGlyIleProPheLeuGlnThrSerAlaIysAsnAlaThrAsnVal 129
Db 466 TTTCGAGACTCTGTGGCATCCCTCTTGGAGAGACGCCAAGAAATGCCCAATGTG 525
Oy 130 GluGlnSerPheMetThrMetAlaAlaGlnIleIysIysArgMetGlyProGlyAlaThr 149
Db 526 GAGCAGGCGTTCATGACCATGCTGCTCTCAAAATCAAAAGCGGATGGGCTGGAGCAGCC 585
Oy 150 AlaGlyAlaGlnIysSerAsnValIysIleGlnSerThrProValIysGlnSerGly 169
Db 586 TCTGGGGGC---GAGCGGCCCAATCTCAAGATCGACACACACCCCTGTAAAGCGGCTGGC 642
Oy 170 GlyIysCysCys 173
Db 643 GTGGCTGTTC 654

RESULT 6
AAI93456
ID AAI93456 standard; cDNA, 1944 BP.
AC
AAI93456;
XX
06-NOV-2001 (first entry)
DE Human polynucleotide SEQ ID NO 13516.
XX
Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KV tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
OS Homo sapiens.
XX
WO200164835-A2.
XX
07-SEP-2001.
XX
26-FEB-2001; 2001WO-US04927.
XX
28-FEB-2000; 2000US-0515126.
XX
18-MAY-2000; 2000US-0577409.
XX
(HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-514838/56.
XX
P-PSDB; AAO13525.
XX
Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
Claim 1; SEQ ID NO 13516; 1399pp + Sequence Listing; English.
XX

```

```

CC The invention relates to human polynucleotides (AAI99941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 1944 BP; 372 A; 618 C; 552 G; 402 T; 0 other.

Alignment Scores:
Pred. No.: 4,72e-78 Length: 1944
Score: 768.50 Matches: 156
Percent Similarity: 80.39% Conservative: 8
Best Local Similarity: 76.47% Mismatches: 7
Query Match: 86.74% Indels: 33
DB: 22 Gaps: 2

US-09-820-003B-2 (1-173) x AAI93456 (1-1944)
Oy 2 SerSerMetAsnProGlnTyrAspTyrLeuPheIysLeuLeuIleGlyAspSerGly 21
Db 56 GCGCCATGAACCCCGAATATGACTTACCTGTTAACTGCTTTGATGGGAGCATCAGGC 115
Oy 22 ValGlyIysSerCysLeuLeuLeuArgPheAlaAspArgTyrThrGluSerTyrIle 41
Db 116 GTGGGCAAGTCATGCTGCTCTCTGCGTTCTATGACAGTACAGAGATACATC 175
Oy 42 SerThrIleGlyValAspPheIysIleArgThrIleGluAsnAspGlyIysThrIleIys 61
Db 176 AGCACCATCGGGGTGAGCTTCAGATCCGAAACCATCGAGTGGAAATATATCAAA 235
Oy 62 LeuGlnIle----- 64
Db 236 CTTGAGATCTGGGACACAGCGGCCAGACGCTTCCGACCATCTCCAGTACTAC 295
Oy 65 -----GluserPheasnAsn 69
Db 296 CGGGGGGCTCATGGCATCTGCTGTATGACGTACTGACCAAGAAATCTTACGCCAAC 355
Oy 70 VallysgIntPLeuGlnIleAspArgTyrAlaSerGluAsnValAsnIysLeu 89
Db 356 GTGAAGCAGTGGCTGACAGAGATTGACCGCTATGCCAGCGAAGCATCAATAACTCTCTG 415
Oy 90 ValGlyAsnIysCysAspLeuThrThrIysIysValValAspTyrThrAlaIysGlu 109
Db 416 GTGGGCAACAGAGAGGAGCTTCAACCAAGAGGTGTGGCAACACACACAGAGAGAG 475
Oy 110 PheAlaAspSerLeuGlyIleProPheLeuGlnThrSerAlaIysAsnAlaThrAsnVal 129
Db 476 TTTCGAGACTCTGTGGCATCCCTCTTGGAGAGAGCGCCAGAAATGCCCAATAGTC 535
Oy 130 GluGlnSerPheMetThrMetAlaAlaGlnIleIysIysArgMetGlyProGlyAlaThr 149
Db 536 GAGCAGGCGTTCATGACCATGCTGCTCAAAATCAAAAGCGGATGGGCTGGAGCAGCC 595
Oy 150 AlaGlyAlaGlnIysSerAsnValIysIleGlnSerThrProValIysGlnSerGly 169
Db 596 TCTGGGGGC---GAGCGGCCCAATCTCAAGATCGACACACACCCCTGTAAAGCGGCTGGC 652
Oy 170 GlyIysCysCys 173
Db 653 GTGGCTGTTC 664

RESULT 7
AAS34668

```

[illegible]

PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488782/53.

DR New polynucleotides and polypeptides for diagnosing, treating,  
 PT preventing or prognosing e.g. diseases or disorders of the nervous,  
 PT musculoskeletal, excretory, gastrointestinal, reproductive, and  
 PT respiratory systems  
 PS  
 PS Disclosure: SEQ ID NO 2092; 642pp; English.

XX The invention relates to novel nucleic acids encoding novel human foetal  
 CC antigens. The nucleic acids and proteins are used to prevent, treat (e.g.  
 CC by gene therapy) or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility  
 CC to a pathological condition. The antibodies to the antigens can also  
 CC be used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. Rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
 CC and ocular disorders e.g. corneal infection. The polypeptides can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, cofactors and other nutritional components. Numerous  
 CC examples of diseases and disorders treated by the nucleic acids and  
 CC proteins are given in the specification. The present sequence  
 CC is a genomic DNA fragment from a gene encoding a foetal antigen of the

Alignment Scores:

Pred. No.: 4,86e-74 Length: 1965  
 Score: 733.50 Matches: 150  
 Percent Similarity: 78.92% Conservative: 11  
 Best Local Similarity: 73.53% Mismatches: 10  
 Query Match: 82.79% Indels: 33  
 DB: 22 Gaps: 2

US-09-820-003B-2 (1-173) x AAS34668 (1-1965)

OY 2 SerSerMetAspProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGly 21  
 Db 68 GCCGCCATGAACCCCGGATGACGCTGTTTAAAGCTGTTTGAATGGGACACAGC 127  
 QY 22 ValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrlSerTyrIle 41  
 Db 128 GTGGGCAAGTCATGCTCTCTCTGCGGTTTCTGCTATGACCTTACACAGAGACTACATC 187

OY 42 SerThrIleGlyValAlaSpPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys 61  
 Db 188 AGCACATCGGGGTGGGATTCAGATCCAAACATCGAGTGGATGGCAAAACTATCAA 247  
 QY 62 LeuGlnIle----- 64  
 Db 248 CTTGAGATCTGGACACGCGCGGCCAGAGAGGTTCTGGACCATCTCCAGCTACTAC 307  
 OY 65 -----GluSerPheAsnAsp 69  
 Db 308 CGGGGGCTCATGGCTTCTCTGCTGATATGACGCTACTGACCAAGAACTCTTATGCCAAC 367  
 QY 70 ValLysGlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeu 89  
 Db 368 GTGAAGCAGTGGCTCAGAGATGATGACCGCATGCCAGCAGAAAGCTCAATTAAGCTCTG 427  
 QY 90 ValGlyAsnLysCysAspLeuThrThrLysLysValAlaAspTyrThrAlaLysGlu 109  
 Db 428 GTGGGCAACAAAGAGGACCTCACCACCAAGAGGTGTGGACACACACAGCCAGAGAG 487  
 QY 110 PheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnVal 129  
 Db 488 TTTCAGACTCTTGGGATCCCTCTTGGAGAGAGAGTGCACAAATGCCAAATATC 547  
 QY 130 GluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThr 149  
 Db 548 GAGCAGCGCTTCATGACCATGCTGATGATCAAAAGCAGATGGGCGCGAGAGACC 607  
 QY 150 AlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerTrpProValLysGlnSerGly 169  
 Db 608 TCCGGGGGC---GAGCGGGCCCAATCTCAAGATCGACAGCACCCCTGTAAAGCGGCTGC 664  
 QY 170 GlyGlyCysCys 173  
 Db 665 GTGGTGTCTTC 676

RESULT 8

ABL05983 standard; cDNA; 1791 BP.

XX ABL05983;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 12431.

XX Drosophila: developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PERE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PsDB; ABB61880.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX Claim 1; SEQ ID NO 12431; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116175-AB130511) expressed DNA sequences (AB101840-AB161275) and the encoded proteins (AB157277-AB172072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/publ/published\\_pcl\\_sequences](http://wipo.int/publ/published_pcl_sequences).

Sequence 1791 BP; 579 A; 379 C; 360 G; 473 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	1791
Score:	2.6e-68	683.00
Percent Similarity:	75.24%	Matches: 143
Best Local Similarity:	69.42%	Conservative: 12
Query Match:	77.09%	Mismatches: 17
		Indels: 34
		Gaps: 3

US-09-820-003b-2 (1-173) x AB105983 (1-1791)

```
OY 1 MetSerMetAspProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
DB 253 ATGTCATCTGTGATTCGGAATACGATATCTTAAAGTGGCTTATGAGACTCG 311
OY 21 GlyValGlySerCysLeuLeuArgPheAlaAspAspThrTyrGluSerTyr 40
DB 312 GCGGTGGAAAGTCGTCCTCTCTGCGGATGCGGACATACATACAAAGATTAT 371
OY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
DB 372 ATCAGCCAAATCGAGTGGATTCTCAAAATCAGACTATAGAACTCGAAGAACATT 431
OY 61 LysLeuGlnIle-----GluSerPheAsn 68
DB 432 AAATCGCAATTCGGAGTACTGTCGGCAGAGCCCTTCCGCAACATCGTCTCATAT 491
OY 65 -----GluSerPheAsn 68
DB 492 TATAGGGCCGCCATGCGATCATCTTCGTATGATTCAGATTCAGCAGCAGAGCTTTCAC 551
OY 69 AsnValLysGlnIlePheLysGlnIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88
DB 552 AATGTGAAGCAGTGGCTGAGAGATCGAGCGTATGCTGTGAGAACTCAACAAGCTG 611
OY 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrAlaLys 108
DB 612 CTGCTGGCAACAGAGCACTTACCCACAGAAAGTAGTCGACACACACAGCTGCG 671
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
DB 672 GAGTACGCCGCCAGTGTAGGCACTTCCTCTGAAACTTGGCCAAAGGCCCCACAC 731
OY 129 ValGlnIleSerPheMetThrMetAlaIleGluIleLysLysArgMetGlyProGlyAla 148
DB 732 GTTAGAGAGAGCCCTCATGATGATGCGGAGATCAACAAATCGCTCGGCCGCGTCC 791
OY 149 ThrAlaGlyLysIleGluLysSerAsnValLysIle---GlnSerThrProValLysIle 167
DB 792 AGCCGCCACTGAC---AACGTCAGCAAGAGTGAAGATCATGACAGAGCTCAGTAGAAAC 848
OY 168 SerGlyLysGlyCysAsp 173
DB 849 ACCAAATCCGCTTCCTGTC 866
```

RESULT 9  
AAFI4053  
ID AAFI4053 standard; cDNA; 1202 BP.  
XX  
AC AAFI4053;

```
XX 13-MAR-2001 (first entry)
DE Aspergillus oryzae EST SEQ ID NO:6576.
DE
XX Multiple gene expression; filamentous fungal cell; EST;
XX Aspergillus oryzae; Fusarium venenatum; Aspergillus niger;
XX Aspergillus oryzae; Trichoderma reesei; identification; recombination;
XX culture condition; environmental stress; spore morphogenesis;
XX metabolic pathway engineering; catabolic pathway engineering; ss.
XX
XX Aspergillus oryzae.
XX
XX WO200056762-A2.
XX
XX 28-SEP-2000.
XX
XX 22-MAR-2000; 2000WO-US07781.
XX
XX 22-MAR-1999; 99US-0273623.
XX
XX (NOVO ) NOVO NORDISK BIOTECH INC.
XX (NOVO ) NOVO NORDISK AS.
XX
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX WPI: 2000-594572/56.
XX
XX Monitoring differential expression of genes in filamentous fungal cells
XX uses fluorescence-labeled nucleic acids isolated from the cells and a
XX substrate of expressed sequence tags -
XX
XX Claim 88: Page 2691; 3161pp; English.
XX
XX The present invention describes a method for monitoring differential
XX expression of genes in a first filamentous fungal (FF) cell relative to
XX expression of the same genes in one or more second filamentous fungal
XX cells. The method uses fluorescence-labeled nucleic acids isolated from
XX the FF cells and a substrate of expressed sequence tags (EST). The ESTs
XX are used in the methods for monitoring differential expression of genes
XX in a first filamentous fungal (FF) cell relative to expression of the
XX same genes in one or more second filamentous fungal cells. Monitoring
XX the global expression of genes from FF cells allows the production
XX potential of the microorganisms to be improved. New genes may be
XX discovered; possible functions of unknown open reading frames can be
XX identified and gene copy number variation and stability can be
XX monitored. The expression of genes can be used to study how FF cells
XX adapt to changes in culture conditions, environmental stress, spore
XX morphogenesis, recombination, metabolic or catabolic pathway
XX engineering. Using ESTs provides several advantages over genomic or
XX random cDNA clones including elimination of redundancy as one spot on an
XX array equals one gene or open reading frame, and organization of the
XX microarrays based on function of the gene products to facilitate
XX analysis of the results. AAF07478 to AAF11247 represents ESTs from
XX Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
XX niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
XX AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
XX all specifically claimed in the present invention.
XX
XX Sequence 1202 BP; 282 A; 294 C; 282 G; 344 T; 0 other;
```

Alignment Scores:

Pred. No.:	Length:	1202
Score:	1.28e-65	657.50
Percent Similarity:	72.82%	Matches: 138
Best Local Similarity:	66.99%	Conservative: 12
Query Match:	74.21%	Mismatches: 19
		Indels: 37
		Gaps: 3

US-09-820-003b-2 (1-173) x AAFI4053 (1-1202)

```
OY 2 SerSerMetAspProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGly 21
DB 111
```





Accession	Gene	Species	Length	Score	Percent Similarity	Best Local Similarity	Query Match	DB
PR 19-JUL-1999	9905-0144334	9905-0159329	14-0CT-1999	9905-0159330	PR 14-0CT-1999	9905-0159331	PR 14-0CT-1999	9905-0159332
PR 20-JUL-1999	9905-0144335	9905-0159333	PR 14-0CT-1999	9905-0159334	PR 14-0CT-1999	9905-0159335	PR 14-0CT-1999	9905-0159336
PR 20-JUL-1999	9905-0144632	9905-0159637	PR 14-0CT-1999	9905-0159638	PR 14-0CT-1999	9905-0159639	PR 14-0CT-1999	9905-0159640
PR 20-JUL-1999	9905-0144884	9905-0159638	PR 14-0CT-1999	9905-0159639	PR 14-0CT-1999	9905-0159640	PR 14-0CT-1999	9905-0159641
PR 21-JUL-1999	9905-0145086	9905-0159584	PR 18-0CT-1999	9905-0159585	PR 18-0CT-1999	9905-0159586	PR 18-0CT-1999	9905-0159587
PR 21-JUL-1999	9905-0145088	9905-0160741	PR 21-0CT-1999	9905-0160742	PR 21-0CT-1999	9905-0160743	PR 21-0CT-1999	9905-0160744
PR 22-JUL-1999	9905-0145085	9905-0160767	PR 21-0CT-1999	9905-0160768	PR 21-0CT-1999	9905-0160769	PR 21-0CT-1999	9905-0160770
PR 22-JUL-1999	9905-0145087	9905-0160768	PR 21-0CT-1999	9905-0160769	PR 21-0CT-1999	9905-0160770	PR 21-0CT-1999	9905-0160771
PR 23-JUL-1999	9905-0145145	9905-0160815	PR 21-0CT-1999	9905-0160816	PR 21-0CT-1999	9905-0160817	PR 21-0CT-1999	9905-0160818
PR 23-JUL-1999	9905-0145218	9905-0160980	PR 22-0CT-1999	9905-0160981	PR 22-0CT-1999	9905-0160982	PR 22-0CT-1999	9905-0160983
PR 26-JUL-1999	9905-0145276	9905-0160989	PR 22-0CT-1999	9905-0160990	PR 22-0CT-1999	9905-0160991	PR 22-0CT-1999	9905-0160992
PR 27-JUL-1999	9905-0145913	9905-0161405	PR 25-0CT-1999	9905-0161406	PR 25-0CT-1999	9905-0161407	PR 25-0CT-1999	9905-0161408
PR 27-JUL-1999	9905-0145918	9905-0161405	PR 25-0CT-1999	9905-0161406	PR 25-0CT-1999	9905-0161407	PR 25-0CT-1999	9905-0161408
PR 28-JUL-1999	9905-0145919	9905-0161359	PR 26-0CT-1999	9905-0161360	PR 26-0CT-1999	9905-0161361	PR 26-0CT-1999	9905-0161362
PR 02-AUG-1999	9905-0146386	9905-0161920	PR 28-0CT-1999	9905-0161921	PR 28-0CT-1999	9905-0161922	PR 28-0CT-1999	9905-0161923
PR 02-AUG-1999	9905-0146388	9905-0161920	PR 28-0CT-1999	9905-0161921	PR 28-0CT-1999	9905-0161922	PR 28-0CT-1999	9905-0161923
PR 03-AUG-1999	9905-0147038	9905-0161923	PR 28-0CT-1999	9905-0161924	PR 28-0CT-1999	9905-0161925	PR 28-0CT-1999	9905-0161926
PR 04-AUG-1999	9905-0147204	9905-0162142	PR 29-0CT-1999	9905-0162143	PR 29-0CT-1999	9905-0162144	PR 29-0CT-1999	9905-0162145
PR 05-AUG-1999	9905-0147302	9905-0162142	PR 29-0CT-1999	9905-0162143	PR 29-0CT-1999	9905-0162144	PR 29-0CT-1999	9905-0162145
PR 05-AUG-1999	9905-0147302	9905-0162142	PR 29-0CT-1999	9905-0162143	PR 29-0CT-1999	9905-0162144	PR 29-0CT-1999	9905-0162145
PR 06-AUG-1999	9905-0147303	9905-0162142	PR 29-0CT-1999	9905-0162143	PR 29-0CT-1999	9905-0162144	PR 29-0CT-1999	9905-0162145
PR 06-AUG-1999	9905-0147416	9905-0162142	PR 29-0CT-1999	9905-0162143	PR 29-0CT-1999	9905-0162144	PR 29-0CT-1999	9905-0162145
PR 09-AUG-1999	9905-0147493	9905-0162142	PR 29-0CT-1999	9905-0162143	PR 29-0CT-1999	9905-0162144	PR 29-0CT-1999	9905-0162145
PR 10-AUG-1999	9905-0148171	9905-0162142	PR 29-0CT-1999	9905-0162143	PR 29-0CT-1999	9905-0162144	PR 29-0CT-199	

DB 714 GCAGAGAGATCTAAGCCACCAACGATCCAGATCCGAGCAACACCGTTAACACGCA--- 770  
Oy 170 GLYGLYCys 173  
DB 771 TCAGGCTCTCTCC 782  
RESULT 11  
AAC35200  
ID AAC35200 standard; DNA; 932 BP.  
AC AAC35200;  
AC  
AC 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9347.  
DE  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140355.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142053.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142927.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145916.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147197.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.



DR WPI: 2001-639362/73.  
DR P-PSDB: ABG19673.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

XX Claim 1; SEQ ID No 19664; 103bp; English.

PS The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1193 BP; 372 A; 279 C; 213 G; 329 T; 0 other;

#### Alignment Scores:

Pred. No.: 8,64e-62 Length: 1193  
Score: 624.00 Matches: 129  
Percent Similarity: 74.57% Conservative: 0  
Best Local Similarity: 74.57% Mismatches: 0  
Query Match: 70.43% Indels: 44  
DB: Gaps: 1

US-09-820-003b-2 (1-173) x AAS83860 (1-1193)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
DB 1003 ATGTCACGACATGATCCGACATATGATTAATTCAGTACTTCTGATGGCAGATCA 944  
QY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspTyrTyrThrGluSerTyr 40  
DB 943 GGGGTGGAAAGCTCTGCTCTTCTTGTAGGTTTGCAGATGATACATATACAGAAAGCTAC 884  
QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
DB 883 ATCAGCAACATTTGGTGTGATTTCAAAATAGAACTATAGAGTTAGACGGAAAAACATC 824  
QY 61 LysLeuGlnIleGluSerPheAsnAsnValLysGlnTrrLeuGlnIleAspArgTyr 80  
DB 823 AACCTCAAAATA----- 812  
QY 81 AlaSerGluAsnValAsnLysLeuLeuValGlyAsnLysCysAspLeuThrThrLysLys 100  
DB 812 ----- 812  
QY 101 ValValAspTyrThrThrAlaLysGluPheAlaAspSerLeuGlyIleProPheLeuGlu 120  
DB 811 -----GAATTTCTGATTCCTTCCCTTGGATTCCTTTTGGAA 776  
QY 121 ThrSerAlaLysAsnAlaThrAsnValGluGlnSerPheMetThrMetAlaAlaGluIle 140  
DB 775 ACCAGTGCTAGATGACAGCAAGATGTAGAAAGCTTTTCAATGACAGATGGCAGCATTT 716  
QY 141 LysLysArgMetGlyProGlyAlaThrAlaGlyLysGluLysSerAsnValLysIle 160

DB 715 AAAAAAGCAATGGGCGGAGCAACACTGTGTGTGAGAAAGTCAATGTTAAATT 656  
QY 161 GlnSerThrProValLysGlnSerGlyValGlyLysCys 173  
DB 655 CAGACACTCCAGTCACAGCACTGAGTGTGAGATTGCTTC 617  
RESULT 13  
AAC42684  
ID AAC42684 standard; DNA; 777 BP.  
AC AAC42684;  
AC 17-OCT-2000 (first entry)  
DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 36471.  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36471.  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
XX 06-SEP-2000.  
PD  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135153.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 21-JUN-1999; 99US-0139763.  
 PR 22-JUN-1999; 99US-0139817.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142390.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143628.  
 PR 15-JUL-1999; 99US-0144003.  
 PR 16-JUL-1999; 99US-0144008.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145228.  
 PR 23-JUL-1999; 99US-0145226.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 22-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 26-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154772.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0155659.  
 PR 29-SEP-1999; 99US-0155659.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158363.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 14-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159337.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:  
 Pred. No.: 1.61e-61  
 Score: 619.50  
 Percent Similarity: 68.97%  
 Best Local Similarity: 61.58%  
 Query Match: 69.92%  
 DB: 21  
 US-09-820-003b-2 (1-173) x AAC42684 (1-777)  
 Length: 777  
 Matches: 125  
 Conservative: 15  
 Mismatches: 30  
 Indels: 33  
 Gaps: 2

```

OY      3 SerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyVal 22
      163 ACCATGAAATCGTACAGACATCTTTCAAGCTCGCTATCGGGATCTGGCGTA 222
OY      23 GlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThiGluSerYrTlleSer 42
      223 GGGCAAGTCTTCTCTTTTGGATGCTCTGATGATGCTTATGTAAGATTACATTTAGC 282
OY      43 ThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLysLeu 62
      283 ACTATTGGAGTGCATTTTAAATATGACTGTGGAACAAGATGGCAAAACATTTAAGCTC 342
OY      63 GluIle----- 64
Db      343 CAATTTGGGACACTGCTGTGTCAGAGACGTTACAGACTATTACTAGACGTTACTACCGT 402
OY      65 -----GluSerPheAsnAsnVal 70
Db      403 GGGGACATCGCAATTTATTTATTTGCTACAGATGTCACAGATGAAGAAAGCTTCATATATGTC 462
OY      71 LysGlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuVal 90
Db      463 AAGCAATGTTGAGTGAATGATGCTTATGCTACTGACAAATGTCACAAACTCTTCTTGT 522
OY      91 GlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThAlaLysGluPhe 110
Db      523 GGAACAAGTGTACTGCTACTGTAACAAAGACGCTCTCTTATGAAAGTCCCAAGCTTTT 582
OY      111 AlaAspSerLeuGlyIleProPheLeuGlnThrSerAlaLysAsnAlaThrAsnValGlu 130
Db      583 GCCCATGAAATCGGATTCCTTTATGAGACTGCTGCAAAAGATGCTACAAACCTGAGAA 642
OY      131 GlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyValAlaThrAla 150
Db      643 CAGGCTTCATGCGCAATGCTGCATCCAAAGAGACAAAGGCTAGCAACACCTGGGG 702
OY      151 GlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGlyGly 170
Db      703 AATAATGAGAACCAACGACCGCTGATGATGAGAGACAGACGCTGTGACAGAG--AAG 759
OY      171 GlyCysCys 173
      760 GGCTGCTGC 768
Db
OY      RESULT 14
      ABN98761
      ID ABN98761 standard; DNA; 881 BP.
      AC ABN98761;
      XX
      DT 01-AUG-2002 (first entry)
      XX
      DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 529.
      XX
      KW Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
      KW disease; crop; thale cress; tolerance factor; insect; pathogen;
      KW nutrition; ds.
      XX
      OS Arabidopsis thaliana.
      OS
      PN US2002023281-A1.
      XX
      PD 21-FEB-2002.
      XX
      PF 26-JAN-2001; 2001US-0770445.
      XX
      PR 27-JAN-2000; 2000US-178472P.
      XX
      PA (GORL/) GORLACH J.
      PA (ANY/) AN Y.
      PA (HAM/) HAMILTON C M.
      PA (PRIC/) PRICE J L.

```

```

PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.
PA (RAME/) RAMEKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHEN A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRICK/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
XX WPI: 2002-400781/43.
XX
XX The invention relates to an Arabidopsis thaliana nucleic acid (I)
XX comprising a sequence capable of hybridizing under stringent conditions
XX to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
XX given in the specification or its fragment. A polypeptide (II) encoded by
XX (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
XX genetically modified cell (IV) comprising an exogenous nucleic acid, is
XX useful for screening a candidate agent for its biological effect. (I) is
XX useful in identifying homologous or related genes, in producing
XX compositions that modulate the expression or function of its encoded
XX protein, mapping functional regions of the protein and in studying
XX associated physiological pathways. (I) is also useful for the genetic
XX manipulation of cells, particularly plant cells. (I) is also useful in
XX screening assays of various plant strains to determine the strains that
XX are best capable of withstanding a particular disease or environmental
XX stress. (II) and (III) are useful for screening of biologically active
XX agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
XX pathways. The screened agents are useful in improved methods of treating
XX crops to prevent or treat disease. (II) are also useful in screening
XX programs to identify agents that mimic or enhance the action of tolerance
XX factors. Such agents are useful in improved methods of treating crops to
XX enhance their tolerance to environmental stress. (I) is also useful
XX for enhancing or inhibiting production of a biosynthetic product in a
XX plant. (III) is useful for identifying other mediators that may induce
XX expression of proteins of interest, for establishing the extent to which
XX any specific insect and/or pathogen is responsible for damage to a
XX particular plant, for identifying other mediators that enhance or induce
XX tolerance to environmental stress, for identifying factors involved in
XX biosynthetic pathways of nutritional, commercial, or medicinal value and
XX for identifying productions of nutritional, commercial or medicinal
XX value. (IV) is useful in the study of genetic function and regulation,
XX for alteration of the cellular metabolism and for screening compounds
XX that may affect the biological function of the gene or gene products.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html?docID=999909770445.
XX
XX Sequence 881 BP; 241 A; 199 C; 180 G; 261 T; 0 other.
XX
XX Alignment Scores:
XX Pred. No.: 1.9e-61 Length: 881
XX Score: 619.50 Matches: 125
XX Percent Similarity: 68.97% Conservative: 15
XX Best Local Similarity: 61.58% Mismatches: 30
XX Query Match: 69.92% Indels: 33

```

[illegible]

```

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144633.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144881.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155653.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

```

```

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

## Alignment Scores:

```

Pred. No.: 9.5e-61      Length: 666
Score: 612.00      Matches: 128
Percent Similarity: 70.56%      Conservative: 11
Best Local Similarity: 64.97%      Mismatches: 24
Query Match: 69.07%      Indels: 34
DB: 21      Gaps: 3

```

US-09-820-003b-2 (1-173) x AAC42764 (1-666)

```

QY 9 Asptyrleuphelsleuleuileglyaspserglyalglylserserleuenu 28
    |||||||
DB 73 GACTATTGTCACAGCTTTCCTCATGCTGATTCGTGGAAAGTCGTGCTT 132
    |||||||

QY 29 Leuargphealaasptyrthrthrgluserthrlieserthrileglyvalasphe 48
    |||||||
DB 133 CTAAAGTTGCTGATGATTCCTACTGATGATGATGATGATGATGATGATGAT 192
    |||||||

QY 49 Lysileargthrilegluleuaspolylsyrthrilelysleuglinile----- 64
    |||||||
DB 193 AAAATCCGACAGTCGACAGATGAGAAAGCATCAACATCCAGATTGGATACGCA 252
    |||||||

QY 64 ----- 64
DB 253 GGCCAGAACGTTCCAGAACGATTACTAGACTTACAGAGAGGCTCATGATATT 312
    |||||||

QY 65 ----- 65
DB 313 CTGACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
    |||||||

QY 77 Ileasparglyralasergluasnvalasnlysleuenuvalglyasnlyscysaspleu 96
    |||||||
DB 373 ATGACGCGCTATGACAGAGGAAAGTTATACAGCTAGCTGTTGGGAACAAGTGTGATCTC 432
    |||||||

QY 97 Thrthrlyslsvalvalasptyrthrthralalysgluphealaaspserserleuglylle 116
    |||||||
DB 433 ACATCACAGAAAGTGTATCCACTGACAGACACTAAGGCTTCCGCTGATGATGATGATGATC 492
    |||||||

QY 117 Prothelenglutirseralalysasnalahrsnvalgluginserphemetthet 136
    |||||||
DB 493 CCATTCTTGAAACAGGCTTAAAGATGCTACCAATGTGCAAGAGCTTTCATGCGCATG 552
    |||||||

```



Oy 137 AlaAlaGluIleLysLysArgMetGlyProGlyAlaThrAlaGlyGlyAlaGlyLysSer 156  
 DB 553 ACTGCTGCATTCAGACGACGAAATGCG--AGCCAACTGCGAGAGATCTAAGCCACCA 609  
 Oy 157 AsnValLysTleGlnSerThrProValLysGlnSerGlyGlyCys 173  
 DB 610 ACGGTCCAGATCCGAGACAACTGTTAACCAACAA---TCAGGCTGCTGC 657

Search completed: June 22, 2003, 18:04:34  
 Job time : 221 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 22, 2003, 16:28:49 ; Search time 1618 Seconds

(without alignments)  
311.734 Million cell updates/sec

Title: US-09-820-003b-2

Perfect score: 886  
Sequence: 1 MSSMNPEDYDFKLLIGDS.....EKSNAKIQSTPVKSGGGCC 173

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame\_plus.model -DRV-xlp  
-Q/cgn2.1/USPFO/050982003/rmat.18062003.144129.13092/app-query.fasta.1.327  
-DB-GenEmbl -QFMT-fastap -SUFFIX-p2n.rge -MINMATCH-0.1 -LOOPEL-0 -LOOPEXT-0  
-UNITs-bits -STAR-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45  
-DOCALLGN-200 -THR.SCOR-pct -THR.MAX-100 -THR.MIN-0 -ALIGN-15 -MODE-LOCAL  
-OUTFMT-pco -NORM-ext -HEAPSIZ-500 -MINLEN-0 -MAXLEN-2000000000  
-USER-050982003.ecgn.1.1.3745-ctunat.18062003.144129.13092 -NCPU-6 -ICPU-3  
-NO\_MMAP -LARGEDUTERY -NDS\_SCORES-0 -WAIT -DSPLOCK-100 -LONOLOG  
-DEV\_TIMEOUT-120 -WARN\_TIMEOUT-30 -THREADS-10 -XGAPOP-6 -XGAPEXT-0.5 -FGAPOP-6  
-FGAPEXT-7 -YGAPEXT-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database :  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pac:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*  
15: gb.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.ju:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pac:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	860	97.1	615	6 E03861	E03861 cDNA encod1
2	860	97.1	618	9 AF498929	AF498929 Homo sapi
3	860	97.1	723	9 HMMRAB1A	M28209 Homo sapien
4	860	97.1	840	4 CRRAB1	X56384 Canine TAB1
5	860	97.1	1428	10 MMRPT1	I00094 Mouse mRNA
6	860	97.1	1433	9 BC009005	BC009005 Homo sapi
7	860	97.1	1444	10 BC002077	BC002077 Mus muscu
8	860	97.1	2686	10 AF226873	AF226873 Mus muscu
9	849	95.8	618	10 NATRASA	J02998 Rat ras-rel
10	838	94.6	841	5 DYGORAB1	M38393 Discopyle o
11	768.5	86.7	925	6 AR070365	AR070365 Sequence
12	768.5	86.7	1985	6 AX086155	AX086155 Sequence
13	768.5	86.7	1985	9 HSM801608	AL136635 Homo sapi
14	762.5	86.1	654	10 RNRAB1B	X13905 Rat cdna fo
15	762.5	86.1	161955	2 AC106368	AC106368 Mus muscu
16	761.5	85.9	1859	10 BC016408	BC016408 Mus muscu
17	758.5	85.6	151988	2 AC123143	AC123143 Rattus no
18	740.5	83.6	1858	2 HSA245875	AF245875 Homo sapi
19	737.5	83.2	185285	2 AC094411	AC094411 Rattus no
20	733.5	82.8	194976	2 AL139414	AL139414 Homo sapi
21	733.5	82.8	194976	2 AL513165	AL513165 Human DNA
22	733.5	82.7	718	3 LSRAB1	AC117841 Rattus no
23	732	81.5	1086	10 RNU63023	X72688 L. stagnalis
24	732	81.5	174028	2 AC099354	AC099354 Rattus no
25	722	79.0	2257	3 AK055927	AK055927 Homo sapi
26	700	77.1	1772	3 DB4312	DB4312 Drosophila
27	683	75.7	952	3 AF013572	AF013572 Bombyx mo
28	671	74.3	720	3 CNS019BR	AF013572 Bombyx mo
29	658.5	74.3	720	3 CNS019BR	AF013572 Bombyx mo
30	655	73.9	1868	3 AY121695	AY121695 Drosophila
31	645	72.8	979	8 LJRAB1C	LJRAB1C
32	644	72.7	938	8 AY052204	AY052204 Arabidops
33	644	72.7	950	8 AF127134	AF127134 Arabidops
34	636.5	71.8	864	8 PEXGRBP10	D12549 Pea mRNA fo
35	634.5	71.6	609	8 AY060495	AY060495 Arabidops
36	633	71.4	636	8 CANS01BW9	AL14769 Bortyflis
37	633	71.4	640	8 AY114013	AY114013 Arabidops
38	633	71.4	931	8 AF324990	AF324990 Arabidops
39	633	71.4	939	8 AY085384	AY085384 Arabidops
40	633	71.4	939	8 AY085384	AY085384 Arabidops
41	632.5	71.4	1008	8 NP433CTP	NP433CTP
42	632.5	71.3	1074	8 AB024994	AB024994 Cicer atl
43	630.5	71.2	1129	8 LJRAB1D	LJRAB1D
44	628.5	70.9	898	8 NTRKBP	X72212 L. tabacum m
45	626.5	70.7	860	8 LJRAB1E	LJRAB1E

## ALIGNMENTS

RESULT 1

E03861 LOCUS E03861 615 bp RNA linear PAT 29-SEP-1997  
 DEFINITION cDNA encoding GTP binding protein yptl.  
 ACCESSION E03861  
 VERSION E03861.1 GI:2172075  
 KEYWORDS JP 1992226000-A/3.  
 SOURCE Mus sp.  
 ORGANISM Mus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 615)  
 Yamauchi, K., Uchida, K., Takai, Y., Aisaka, K. and Sakurada, K.  
 LOW MOLECULAR WEIGHT PROTEIN BINDING TO GUANOSINE TRIPHOSPHATE  
 Patent: JP 1992226000-A 3 14-AUG-1992;  
 KYOMA HAKKO KOGYO CO LTD  
 COMMENT  
 OS Mus sp. (mouse)  
 PN JP 1992226000-A/3  
 PD 14-APR-1992  
 PF 25-APR-1991 JP 1991095304  
 PR 21-MAY-1990 JP 90P 130547  
 PI YAMAGUCHI KAZUO, UCHIDA KAZUHISA, TAKAI YOSHIMI, AISAKA KAZUO,  
 PI SAKURADA KAZUHIRO  
 PC C07K15/18,C07K13/00,C07K15/06,C12N15/12,G01N33/50//A61K37/02;  
 CC strandedness: Single;  
 topology: Linear.  
 FEATURES  
 source Location/Qualifiers  
 1..615  
 /organism="Mus sp."  
 /db\_xref="taxon:10095"  
 BASE COUNT 196 a 117 c 153 g 149 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.51e-75 Length: 615  
 Score: 860.00 Matches: 173  
 Percent Similarity: 84.39% Conservative: 0  
 Best Local Similarity: 84.39% Mismatches: 0  
 Query Match: 97.07% Indels: 32  
 DB: Gaps: 1  
 US-09-820-003b-2 (1-173) x E03861 (1-615)  
 QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20  
 DB 1 ATGTCACAGATGAATCCGCAATATGATTTATTCAGTACTTCTGATGGCGATTC 60  
 QY 21 GlyValIGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40  
 DB 61 GGGGTGGAAATCCTGCTTCTCTAGAGTTTGCAGATGATACGATACGGAAGCTAC 120  
 QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyTyrThrIle 60  
 DB 121 ATCAGCACAAATGTGTGATTTCAAGATACGACATATGATGATGGAAACCAATC 180  
 QY 61 LysLeuGlnIle----- 64  
 DB 181 AAGCTACAGATATGGGACACAGCAGCCAGGAAGATTTGCAACATCACTTCCAGTTAT 240  
 QY 65 -----GluSerPheasn 68  
 DB 241 TACAGAGAGAGCCATGGCATCATAGTTGTGTATGATGTGACAGATCAGAGATCCTTCAT 300  
 QY 69 AsnValLysGlnTTPLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88  
 DB 301 AACGTTAAACAGTGTGCTCAGAGATAGATCCCTACGCGAGGAAATGTCACAAAGTTG 360  
 QY 89 LeuValIGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108  
 DB 361 TTGGTAGGAAACAAATGTGACTGACACACAAAGAAAGTAGTACTACACAAACAGCAAG 420  
 QY 109 GluPheAlaAspSerLeuGlyTyrLeuPheLeuGluTyrSerAlaLysAsnAlaThrAsn 128  
 DB 421 GAATTTGACGATTCCTGCTGATTCATTTTGGAAACCAAGTCTAAGACCAACAGAT 480

QY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148  
 DB 481 GTAGAACACTCTTTCATGACGATGACGAGCTGAGATTAATAAGCAATGGCTCGAGCT 540  
 QY 149 ThrAlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168  
 DB 541 ACAGCTGTGTGGCCGAGAAATGTCATGTAAATCCAGACACTCCAGTCAAGCAGTCA 600  
 QY 169 GlyGlyGlyCysCys 173  
 DB 601 GGTGAGAGCTGCTGC 615  
 RESULT 2  
 AF498929 618 bp mRNA linear PRI 01-MAY-2002  
 LOCUS AF498929  
 DEFINITION Homo sapiens small GTP binding protein RAB1A (RAB1A) mRNA, complete  
 cds  
 ACCESSION AF498929  
 VERSION AF498929.1 GI:20379033  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 618)  
 Puhl, H.L., Ili, Ikeda, S.R. and Aronstam, R.S.  
 Homo sapiens RAB family small GTP binding protein RAB1A  
 Unpublished  
 2 (bases 1 to 618)  
 Puhl, H.L., Ili, Ikeda, S.R. and Aronstam, R.S.  
 Direct Submission  
 Submitted (05-APR-2002) CDNA Resource Center, Guthrie Research  
 Institute, One Guthrie Square, Sayre, PA 18840, USA  
 FEATURES  
 source Location/Qualifiers  
 1..618  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /tissue\_type="Brain"  
 1..618  
 /gene="RAB1A"  
 /gene="RAB1A"  
 /note="RAB1A"  
 /note="RAB family member"  
 /codon\_start=1  
 /product="small GTP binding protein RAB1A"  
 /protein\_id="AA021077.1"  
 /db\_xref="GI:20379034"  
 /translation="MSSMPEDYDFKLLIDSGVGSCLLIRADPTYESYIST  
 GDFKIRITLEDGKTKLQIMDTAQERRTTSSYRGAGIIVYVDTQSENNY  
 KQMLQIDRYASENVNKLIVGKCKDLTKKVDYTAKEAFADSLGIPLETSAKNATN  
 VQSEFTMAAEIKRMKRGPAATGAEKSVKIOSIPVQSGGCG"  
 BASE COUNT 203 a 110 c 149 g 156 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.53e-75 Length: 618  
 Score: 860.00 Matches: 173  
 Percent Similarity: 84.39% Conservative: 0  
 Best Local Similarity: 84.39% Mismatches: 0  
 Query Match: 97.07% Indels: 32  
 DB: Gaps: 1  
 US-09-820-003b-2 (1-173) x AF498929 (1-618)  
 QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20  
 DB 1 ATGTCACAGATGAATCCGCAATATGATTTATTCAGTACTTCTGATGGCGACTCA 60  
 QY 21 GlyValIGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40  
 DB 61 GGGGTGGAAATCCTGCTTCTCTAGAGTTTGCAGATGATACGATACGGAAGCTAC 120

```

OY 41 IleserThrilleglyValaAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
DB 121 ATCAGACCAATGTCGTGATTCACAAATATAGACTATAGAGTCGCGGAAAAATCATC 180
OY 61 LysLeuGlnIle----- 64
DB 181 AGCTTCATATATGAGACACAGACGCCAGAAAGATTTCACATACCTCCAGTTAT 240
OY 65 -----GluSerPheAsn 68
DB 241 TACAGAGACCCATGCGATCATAGTTCGTATGATGACAGATCAGGAGTCCTTCAT 300
OY 69 AsnValIysGlnIlePheGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88
DB 301 AATGTTAAACAGTGGCTGAGAAATAGATCGTTATGCCAGTGAATAATGTCACAAATATG 360
OY 89 LeuValIleAsnLysCysAspLeuThrThrLysValValaAspTyrThrAlaLys 108
DB 361 TTGGTAGGAGCAAAATGATCTGACCAAAAGAAAGTAGAGACTACACACAGCGAAG 420
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
DB 421 GAATTCGCTGATCCCTTGGAATTCGCTTTTGGAAACCACTGCTAAGAAATCAAGAT 480
OY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
DB 481 GTAGACACAGCTTTCATGACATGCGACGCTGAGATTTAAAAAGCAATGGCTCCGAGACA 540
OY 149 ThrAlaGlyIleAlaGluLysSerAsnValIysIleGlnSerThrProValIysGlnSer 168
DB 541 ACAGCTGGTGGCTTCAGAGATCCAAATGTTAAATTCAGACACTCCAGTCAAGACGTCA 600
OY 169 GlyIleGlyCysCys 173
DB 601 GGTGGAGGTTGCTGC 615

RESULT 3
HOMERBLA 723 bp mRNA linear PRI 08-JAN-1995
DEFINITION Homo sapiens GTP-binding protein (RAB1) mRNA, complete cds.
ACCESSION M28209.1 GI:550059
VERSION M28209.1 GI:550059
KEYWORDS GTP-binding protein; ras oncogene.
SOURCE Homo sapiens (tissue library: of J.Mallet) pheochromocytoma cdna to mRNA.
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  1 (bases 1 to 723)
  Zahirzouli A, Touchot N., Chardin P. and Tavitian A.
  The human Rab genes encode a family of GTP-binding proteins related
  to yeast YPL1 and Sec4, products involved in secretion
  J Biol. Chem. 264 (21), 12394-12401 (1989)
  JOURNAL 89308668
  MEDLINE 2301306
  COMMENT On Sep 27, 1994 this sequence version replaced gi:341862.
  FEATURES
    Location/Qualifiers
      1..723
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /map="Unassigned"
        /tissue_type="pheochromocytoma"
        /tissue_lib="Of J.Mallet"
        1..723
        /gene="RAB1"
        /gene="RAB1"
        51..668
        /codon_start=1
        /product="GTP-binding protein"
        /protein_id="AA60240.1"
        /db_xref="GI:550060"
        /db_xref="GB:G00-118-857"
        /translation="MSSMNPEDYLFKLLIGDSGVKSCLLRPADDTYESTIST"
  gene
  CDS

```

```

BASP COUNT 225 a 142 c 180 g 176 t
ORIGIN
Alignment Scores:
Pred. No.: 5,46e-75 Length: 723
Score: 860.00 Matches: 173
Percent Similarity: 84.39% Conservative: 0
Best Local Similarity: 84.39% Mismatches: 0
Query Match: 97.07% Indels: 32
DB: 9 Gaps: 1

US-09-820-003b-2 (1-173) x HOMERBLA (1-723)
OY 1 MetSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
DB 51 ATGTCGACATGATCCCGAAATATGATTTATTCAGTTACTTCGATTCGAGACTCA 110
OY 21 GlyValIleLysSerCysIleuLeuAspPheAlaAspAspThrTyrThrGluSerTyr 40
DB 111 CGGTTGGAAATCTTCGCTTCCTTCCTTACGTTTCACATATACATACGAAAGCTAC 170
OY 41 IleserThrilleglyValaAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
DB 171 ATCAGACCAATGTCGTGATTCACAAATATAGACTATAGAGTCGCGGAAAAATCATC 230
OY 61 LysLeuGlnIle----- 64
DB 221 AATGTTAAACAGTGGCTGAGAAATAGATCGTTATGCCAGTGAATAATGTCACAAATATG 290
OY 65 -----GluSerPheAsn 68
DB 291 TACAGAGACCCATGCGATCATAGTTCGTATGATGTCAGACATCAGGATCCCTTCAT 350
OY 69 AsnValIysGlnIlePheGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88
DB 351 AATGTTAAACAGTGGCTGCGCAAGAAATAGATCGTTATGCCAGTGAATAATGTCACAAATATG 410
OY 89 LeuValIleAsnLysCysAspLeuThrThrLysValValaAspTyrThrThrAlaLys 108
DB 411 TTGGTGGGAAATATGATGATCGACACAAAGAAATAGTACTACACACAGCGAAG 470
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
DB 471 GAATTCGCTGATCCCTTGGAATTCGCTTTTGGAAACCACTGCTAAGAAATGTCACAAATATG 530
OY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
DB 531 GTAGACACAGCTTTCATGACATGCGACGCTGAGATTTAAAAAGCAATGGCTCCGAGACA 590
OY 149 ThrAlaGlyIleAlaGluLysSerAsnValIysIleGlnSerThrProValIysGlnSer 168
DB 591 ACAGCTGGTGGCTTCAGAGATCCAAATGTTAAATTCAGACACTCCAGTCAAGACAGTCA 650
OY 169 GlyIleGlyCysCys 173
DB 651 GGTGGAGGTTGCTGC 665

RESULT 4
CFRAB1 840 bp mRNA linear MAM 30-SEP-1999
DEFINITION Canine rab1 mRNA for ras-related GTP-binding protein involved in membrane traffic.
ACCESSION X56384
VERSION X56384.1 GI:913
KEYWORDS GTP-binding protein; protein transport; ras gene; ras gene family.
SOURCE Canis familiaris.
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
  1 (bases 1 to 840)
  REFERENCE

```

AUTHORS Chavrier, P., Vigneron, M., Sander, C., Simons, K. and Zerial, M.  
 TITLE Molecular cloning of YP1/SEC4-related cDNAs from an epithelial cell line  
 JOURNAL Mol. Cell. Biol. 10 (12), 6578-6585 (1990)  
 MEDLINE 91061765  
 PUBMED 2123294  
 REFERENCE 2 (bases 1 to 840)  
 AUTHORS Zerial, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-AUG-1990) Zerial M., EMBL, Meyerhofstrasse 1, 6900 Heidelberg, FRG

FEATURES  
 source Location/Qualifiers  
 1..840  
 /organism="Canis familiaris"  
 /strain="cocker spaniel"  
 /db\_xref="taxon:9615"  
 /clone="47"  
 /cell\_line="MDCK"  
 149..766  
 /gene="rab1"  
 149..766  
 /gene="rab1"  
 /function="GTP binding protein"  
 /codon\_start=1  
 /protein\_id="CAB56775.1"  
 /db\_xref="GI:6006435"  
 /translation="MSSMNPEDYLFKLLIGDSGVKSCLLRFPADPTYESYISTI  
 GVDFKIRTELDGKTKIOLIMPTAGOEFRITTSYRGAGIIVYVDTOESNNV  
 KMWLEIDRYASENVNKLIVGKCDLTKRKVDYTTAKFAPASITIIPILETSAKNATN  
 VQOSPTMAAEIKRMRGPRATAGAGAEKSNVAKIQSTIPVQSGGCC"

BASE COUNT 250 a 176 c 220 g 194 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 6.52e-75 Length: 840  
 Score: 860.00 Matches: 173  
 Percent Similarity: 84.39% Conservative: 0  
 Best Local Similarity: 84.39% Mismatches: 0  
 Query Match: 97.07% Indels: 32  
 DB: 4 Gaps: 1

US-09-820-003B-2 (1-173) x CFRAB1 (1-840)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20  
 DB 149 ATGTCACGACATGAATCCGAATATGATATTATTCACACTTCTGATGGCGACTCT 208  
 QY 21 GlyValGlyLysSerCysLeuLeuLeuAlaGlyPheAlaAspAspThrTyrThrGluSerTyr 40  
 DB 209 GGGGTTGGAAAGCTTGGCTCTCTTCTAGGTTGCAGATGATACATATACAGAAAGCTAC 268  
 QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
 DB 269 ATCAGACACATGGTGCTGATTTCAAAATTAAGACTATAGAGTTAGATGGAAACAAATC 328  
 QY 61 LysLeuGlnIle----- 64  
 DB 329 AAGCTTCAAAATATGGACACACAGCGCCAAAGAAAGATTTCGAACAATCAGCTCCAGTTAT 388  
 QY 65 -----GluSerPheAsn 68  
 DB 389 TACAGAGGAGCCCATGGCATCATAGTGTGTATGACGTGACAGATCAGAGAGCTCTCAAT 448  
 QY 69 AsnValLysGlnTyrLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88  
 DB 449 AATGTTAAACATGGCTGACGAAATAGACCGCTTATGCCAGTGAAGAAAGCTCAACAGATTG 508  
 QY 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrAlaLys 108  
 DB 509 TTGGTAGGAGCAAAATGCGATCTGACCAAGAAAGTAGTAGACTACACAGCAAG 568  
 QY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128

DB 569 GAATTGGCGGATTCCTTGGAAATCCATTTTTGGAAACCAAGTCTAGAACGCCAAGAT 628  
 QY 129 ValGluGlnSerPheMetThrMetAlaIleGluIleLysArgMetGlyProGlyAla 148  
 DB 629 GTAGAACACTCTTTCAATGACGATGCGAGCTGATTAAGAAAGCAATGGGCTCTGGACA 688  
 QY 149 ThrAlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168  
 DB 689 ACAGCTGGTGGTGCAGAGAGATGCCAATGTGAAATTCAGAGCAGCTCCGGTCAAGCAGTCA 748  
 QY 169 GlyGlyGlyCysCys 173  
 DB 749 GGTGGAGGTTCCTGC 763

RESULT 5  
 MMYPT1 1428 bp mRNA linear ROD 12-SEP-1993  
 LOCUS MMYPT1  
 DEFINITION Mouse mRNA for ras-related YP1 protein.  
 ACCESSION Y00094.1 GI:55456  
 VERSION Y00094.1 GI:55456  
 KEYWORDS ras-related protein; YP1 gene.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCES  
 1 (bases 1 to 1428)  
 Gallwitz, D.  
 Direct Submission  
 Submitted (13-NOV-1987) Gallwitz D., Max-Planck-Institut fuer Biophysikalische Chemie, Abt. Molekulare Genetik, Postfach 2841, D-3400 Goettingen, FRG  
 2 (bases 1 to 1428)  
 Haubruck, H., Disela, C., Wagner, P. and Gallwitz, D.  
 The ras-related YP1 protein is an ubiquitous eukaryotic protein: isolation and sequence analysis of mouse cDNA clones highly homologous to the yeast YP1 gene  
 EMBO J. 6 (1987) In press  
 JOURNAL Location/Qualifiers  
 1..1428  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_line="F9 and C3H10T1/2"  
 130..747  
 /note="YP1 protein (AA 1-205)"  
 /codon\_start=1  
 /protein\_id="CAA68284.1"  
 /db\_xref="GI:55457"  
 /db\_xref="MGD:MG1:97842"  
 /db\_xref="SWISS-PROT:P11476"  
 /translation="MSSMNPEDYLFKLLIGDSGVKSCLLRFPADPTYESYISTI  
 GVDFKIRTELDGKTKIOLIMPTAGOEFRITTSYRGAGIIVYVDTOESNNV  
 KMWLEIDRYASENVNKLIVGKCDLTKRKVDYTTAKFAPASITIIPILETSAKNATN  
 VQOSPTMAAEIKRMRGPRATAGAGAEKSNVAKIQSTIPVQSGGCC"

misc.feature  
 polyA\_site  
 1428  
 /note="pot. polyA signal"  
 BASE COUNT 410 a 263 c 324 g 431 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.22e-74 Length: 1428  
 Score: 860.00 Matches: 173  
 Percent Similarity: 84.39% Conservative: 0  
 Best Local Similarity: 84.39% Mismatches: 0  
 Query Match: 97.07% Indels: 32  
 DB: 10 Gaps: 1

US-09-820-003B-2 (1-173) x MMYPT1 (1-1428)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20  
 DB 130 ATGTCACGACATGAATCCGAATATGATATTATTCACACTTCTGATGGCGACTCT 189



[illegible]

Pred. No.:	1,24e-74	Length:	1444
Score:	860.00	Matches:	173
Percent Similarity:	84.39%	Conservative:	0
Best Local Similarity:	84.39%	Mismatch:	0
Query Match:	97.07%	Indels:	32
DB:	10	Gaps:	1

  

US-09-820-003B-2 (1-173) x BC002077 (1-1444)	
QY	1 MetSerSerMetAsnProGluTyrAspPyrLeuPheLysLeuLeuIleGlyAspSer 20
Db	196 ATGTCCAGCAGTAAATCCCGAATATGATTATTATTATTCAGTACTTCGATGGGAGTCT 255
QY	21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspPheTyrThrGluSerTyr 40
Db	256 GGGGTTGGAAAGTCTCTGCTCTCTCTCTTATAGTTTCAGATGATAGTATACGGAAGCTAC 315
QY	41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
Db	316 ATCGACCAATATGGTGTGATTCACAGTATCAGAACTATAGATTAGATGGGAAACAATC 375
QY	61 LysLeuGlnIle----- 64
Db	376 AAGCTACAGATATGGGACACAGACGAGCCGAGAAAGATTGCAACAATCACTCCAGTTAT 435
QY	65 -----GluSerPheAsn 68
Db	436 TACAGAGAGCCCATGCGATCATGTTGTGTATGATGACAGATACAGATGAGAGTCCCTCAAT 495
QY	69 AsnValLysGlnTyrPheLysGlnIleLysAspArgTyrIleSerGluAsnValAsnLysLeu 88
Db	496 AACCTTAAACAGTGGCTGACGAGATGATGCTACGCGACGTGAAATGTCACCAAGCTTG 555
QY	89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108
Db	556 TTGTTAGAGGAACAATGTGACCTGACCCACAAGAAAGTAGTAGACTACACAAACGAGAAAG 615
QY	109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
Db	616 GAATTTTCAGATTCCTCTTGAATTCACATTTTGGAAACCACTGCTAAGAACGCAACGAAAT 675
QY	129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyVala 148
Db	676 GTAAACAGCTTTCATGACAGATGACAGCTAGATTTAAAAAGCGAATGGCTCTGGAGCT 735
QY	149 ThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
Db	736 ACACCTGTGTGTCGCGAGAAAGTCCAAAGTTTAAATCCAGACACTCCAGTCACAGACATCA 795
QY	169 GlyGlyGlyCysCys 173
Db	796 GGTGGAGGCTGCTGCTGC 810

  

RESULT 8	AF226873	2686 bp	mRNA	linear	ROD 13-FEB-2000
LOCUS	AF226873				
DEFINITION	Mus musculus small GTP-binding protein RAB1A mRNA, complete cds.				
ACCESSION	AF226873				
VERSION	AF226873.1				
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					



FEATURES	Location/Qualifiers
SOURCE	1 3686

```

/organism="Mus musculus"
/db_xref="taxon:10090"
29.646
CDS

```

BASE COUNT 778 a 520 c 563 g 825 t

QY 109 GIUPHEALASPSELEUGLYLLEPROPELEUGLUTHRSERALALYSANALATHRASN 128  
 DB 421 GAATTTGACGATTCCTTGGAATTCATTTTGGAAACAGCTGTAAGCAAAAAGAAAT 480  
 QY 129 VALGIUGINSERPHEMETHERMETAALAGLUILELYLSYASRGMEGLPROGLYALA 148  
 DB 481 GAGAACACTCTTTCATGACCATGCGAGCGAGATTAAACCGCATGGCTCTGAGACA 540  
 QY 149 THRAGLGLYALAGLULYSERANVALYLSLEGINSERTHPROVALLYSGLSER 168  
 DB 541 ACAGCTGAGCTGCGAGAGATCCATGTTAAATCCAGACACATCCAGCTAAGACGTA 600  
 QY 169 GLYGLYGLYCYS 173  
 DB 601 GGTGAGGCTGCTGC 615

RESULT 10  
 DYGORAB1 841 bp mRNA linear VRF 28-APR-1993  
 LOCUS DISCOPYGE ommata GTP-binding protein (o-rab1) mRNA, complete cds.  
 DEFINITION M38393.1 GI:213122  
 VERSION GTP-binding protein.  
 KEYWORDS D.ommata electric lobe electromotor neuron, cDNA to mRNA.  
 SOURCE Discopyge ommata  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; Elasmobranchii; Squalae; Hynnosqualae; Pristiogorae; Batoidae; Torpediniformes; Narcinoidae; Narcinidae; Discopyge.  
 REFERENCE Ngsee,J.K., Eiferink,L.A. and Schellier,R.H.  
 A family of ras-like GTP-binding proteins expressed in electromotor neurons  
 Unpublished (1990)  
 JOURNAL Draft entry and computer-readable sequence for [1] kindly submitted by J.K.Ngsee, 10-SEP-1990.  
 COMMENT Dept of Biological Sciences  
 Herrin lab  
 Stanford University  
 Stanford, CA 94305-5020  
 Draft entry and computer-readable sequence for [1] kindly submitted by J.K.Ngsee, 10-SEP-1990.  
 DEPT OF BIOLOGICAL SCIENCES  
 Herrin lab  
 Stanford University  
 Stanford, CA 94305-5020.  
 FEATURES  
 source  
 1..841  
 /organism="Discopyge ommata"  
 /db\_xref="taxon:7785"  
 /cell\_type="electromotor neuron"  
 /tissue\_type="electric lobe"  
 1..841  
 /gene="o-rab1"  
 79..687  
 /gene="o-rab1"  
 /codon\_start=1  
 /product="GTP-binding protein"  
 /protein\_id="AAA49234.1"  
 /db\_xref="GI:213123"  
 /translation="MNPEYDYLKLLIGDSGVGSCLLRPADDTYTESYITGVNFKIRIELDKTKILQIMDTGQGERFRITSYRGAGIITIVYDVDTQESFNKKQVLOEIDRYSERVNKLIVGNKCDLTTKKVVDTTAKERADSLGIPLETSAKNATVEQAFMTMAEIKRMDPGATSGSGSEKSNVNIQSTPKVSGSGGCC"  
 BASE COUNT 269 a 154 c 194 g 224 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9.05e-73 Length: 841  
 Score: 838.00 Matches: 167  
 Percent Similarity: 83.41% Conservative: 4  
 Best Local Similarity: 81.46% Mismatches: 2  
 Query Match: 94.58% Indels: 32

DB: 5 Gaps: 1  
 US-09-820-003b-2 (1-173) x DYGORAB1 (1-841)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20  
 DB 70 ATGTCACCAATGAAATCCGAAATATGACTATTATTCAGAGCTGCTGTTATGGACATCA 129  
 QY 21 GlyValIglyLysSerCysLeuLeuLeuArgPheAlaAspThrTyrThrGluSerTyr 40  
 DB 130 GGTGTGGAAATCTTGCTCTCTCTCTAGATTGACATGATACATATACAGAAAGTAT 189  
 QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60  
 DB 190 ATCAGTACAAATGCTGTGATTTTAAATCAGAACAAATAGATTAGACGCAAGCATTTC 249  
 QY 61 LysLeuGlnIle----- 64  
 DB 250 AAACCTCAAAATCTGGGACACGGCTGTCAGAGCGGTTTCGAACAATCAGCTCCAGTTAC 309  
 QY 65 -----GluSerPheasn 68  
 DB 310 TACAGAGGGGCACATGGCATCATAGTTGTATGATGTTACAGACCAGAGCTGTTTAC 369  
 QY 69 AsnValLysGluTTPLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88  
 DB 370 AATGTAAACATAGCTTCAGGAATATGATGCTTATGCAAGTAAATGTTAACAGTTA 429  
 QY 89 LeuValIglyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108  
 DB 430 CTGGTGGGAAACAAATGTGATCTGCACACAAAGAGGTGGATTTATACAAAGCAAG 489  
 QY 109 GIUPHEALASPSELEUGLYLLEPROPELEUGLUTHRSERALALYSANALATHRASN 128  
 DB 490 GAATTTGACGACTCCCTTGATCCCATTTTGGAAACAGTGCAGAAATGCAACAAAT 549  
 QY 129 VALGIUGINSERPHEMETHERMETAALAGLUILELYLSYASRGMEGLPROGLYALA 148  
 DB 550 GTAGAACCAAGCTTCATGACTATGCTCTCTAGATTAAACGAAATGGCTCGGAGCT 609  
 QY 149 THRAGLGLYALAGLULYSERANVALYLSLEGINSERTHPROVALLYSGLSER 168  
 DB 610 ACATCTGTGCTGCTCGAAAGCAATTAATATATACAAACAGCCAGTGAATGCTCT 669  
 QY 169 GLYGLYGLYCYS 173  
 DB 670 GGTGAGGCTGCTGC 684

RESULT 11  
 AR070365 925 bp DNA linear PAT 18-FEB-2000  
 LOCUS Sequence 4 from patent US 5892012.  
 DEFINITION AR070365  
 ACCESSION AR070365  
 VERSION AR070365.1 GI:7221253  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 925)  
 Hillman,J.L., Lal,P., Corley,N.C. and Shah,P.  
 Rab proteins  
 Patent: US 5892012-A 4 06-APR-1999;  
 JOURNAL Location/Qualifiers  
 TITLE 1..925  
 ORIGIN

BASE COUNT 198 a 263 c 288 g 176 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.92e-66 Length: 925  
 Score: 768.50 Matches: 156  
 Percent Similarity: 80.39% Conservative: 8  
 Best Local Similarity: 76.47% Mismatches: 7

[illegible]

Pred. No.:	146-65	Length:	1985
Score:	768.50	Matches:	156
Percent Similarity:	80.39%	Conservative:	7
Best Local Similarity:	76.47%	Mismatches:	8
Query Match:	86.74%	Indels:	33
DB:	6	Gaps:	2

  

US-09-820-003B-2 (1-173) x AX086155 (1-1985)	
OY	2 SerSerMetasnpProglutryAspTyrleuPheLysleuLeuIleGlyAspSergly 21
Db	42 GCGCCATCAACCCCGAATATGACATCCCTTTTAAGCTGCTTTGATTGGCACTACGCG 101
OY	22 ValGlyLysSercysleuLeuLeuArgPheAlaAspThrTyrThrGluSeryTyrle 41
Db	102 GTGGGGCAATCATACCTCGCTCCCTCGCGCTTGGCTGATGACACGTACACAGAAAGCTCATAC 161
OY	42 SerThrIleGlyValAlaPheLysIleArgTyrThrIleGluLeuAspGlyLysThrIleLys 61
Db	162 AGCCACATCGGGGGGTGGACTTCAAAGATCGAACCATCGAGCTGATGGCAAAATCATCAAA 221
OY	62 LeuAlaIle----- 64
Db	222 CTCACATCTGGGACACAGCGGGCCAGGAAGCTTCGGACCATCACTTCGAGCTACTAC 281
OY	65 -----GluSerPheAsnAsn 69
Db	282 CGGGGGGCTCATGCGATCATCGTGGTGTATGACGTACTGACACAGAAATCTCCAGCGCAAC 341
OY	70 ValLysGlnTyrPheLysGlnIleAspArgTyrAlaSergLysAsnValAlaAspLysleu 89
Db	342 GTGAACCATGTGGCTGACGAGATATGACGGTATAGCCAGACGAAACGTCAATTAAGCTCTG 401
OY	90 ValGlyAsnLysCysAspLeuThrThrLysLysValAlaAspTyrThrThrAlaLysGlu 109
Db	402 GTGGGCAAAAGACGACCTCTACACCAAGAAAGTGTGGACCAACACCAAGCCAAAGAG 461
OY	110 PheAlaAspSeriLeuGlyIleProPheLeuGlnThrSerAlaLysAsnAlaThrAsnVal 129
Db	462 TTTCGACATCTCTGGGACATCCCTCTTGTGGACAGACGCGCAAGATCCCAACCATATGTC 521
OY	130 GluGlnSerPheMetThrMetAlaAlaGluLysLysArgMetGlyProGlyAlaThr 149
Db	522 GAGCAGAGCCTTCAATGACCAATGAGCTCTGTAATATAAAAGCGAAGGGGCGCTGGAGGCC 581
OY	150 AlaGlyLysAlaGluLysSersAsnValLysIleGlnSerThrProValLysGlnSergly 169
Db	582 TCTGTGGGGGG- -GAGCGCGCCCAATCTCAAGATGACAGCACCCTGTAAAGCGCGGTGCG 638
OY	170 GlyGlyCysCys 173
Db	639 GGTGGCTGTTCG 650
RESULT 13	
LOCUS	HSMB01608 1985 bp mRNA linear PRI 20-MAR-2002
DEFINITION	Homo sapiens mRNA; cDNA DKFP564I172 (from clone DKFP564I172);
ACCESSION	complete cds.
VERSION	AL136635
KEYWORDS	AL136635.1 GI:12052795
SOURCE	human
ORGANISM	Homo sapiens
REFERENCE	Bukacinska; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1985)
AUTHORS	Duesterhoeft,A., Lauber,J., Newes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152 Muenster, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;



```

OY 2 SerSerMetAsnProGluTyrAspTyrLeuPheLeuLeuLeuIleGlyAspSergly 21
DB 6 GCCCCATGACACCCGATATGACACTGCTTAAAGCTGCTTTGATGTGGTACCTGGCG 65
OY 22 ValAllylSerCysLeuLeuLeuLeuLeuPheAlaAspPheThrGlyThrGlySerTyrIle 41
DB 66 GTGGGCAAGCATGCGCTCTCTGCGGTTTGTGTATGACACGATGACAGAGATGATATC 125
OY 42 SerThrIleGlyAlaAspPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 61
DB 126 AGCACCTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 185
OY 62 LeuGlnIle----- 64
DB 186 CTACAGATTGGGACACAGCTGCTGACAGAACGCTTCAGACACTGCTGACCTACTAT 245
OY 65 -----GluSerPheAsnAsn 69
DB 246 CGGGGTGCTCAGCAGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
OY 70 ValIleGlnIlePheGlnIleAlaAspArgTyrAlaSerGluAsnValAlaAspLeuLeu 89
DB 306 GTGAAACAGTGGCTGCTGAGAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 365
OY 90 ValGlyAsnLysCysAspLeuThrThrThrLysValValAlaAspTyrThrAlaLysGlu 109
DB 366 GTACGCAACAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
OY 110 PheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnVal 129
DB 426 TTTCACACACTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 485
OY 130 GluIleAspPheMetThrMetAlaAlaGluIleLysLysArgMetCylProGlyAlaThr 149
DB 486 GAAAGGCAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 545
OY 150 AlaGlyGlyAlaGlyLysSerAsnValLysIleGlnIleSerProValLysGlnIleSer 169
DB 546 TCTGGGGGT---GAAGGGCCCAACCTGAGATGACGACGACACTGCTGTAATCTGTAG 602
OY 170 GlyGlyCysArg 173
DB 603 GGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 614

RESULT 15
AC106368/c 161955 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-117N16, *** SEQUENCING IN PROGRESS
DEFINITION *** 71 unordered pieces.
ACCESSION AC106368.2 GI:21731746
VERSION AC106368.2
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
          Chordata; Craniata; Vertebrata; Euteleostomi;
          Eukaryota; Metazoa; Chordata; Sclerozoa; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 161955)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
        Albrooke,S.L., Amaralunge,H.C., Are,J.R., Ayala,M., Banks,T.,
        Barbato,J., Benton,J., Blumhagen,K., Blankenburg,K., Bonni,D.,
        Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
        Bunay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
        Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
        Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
        Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
        Davila,M.L., Davis,C., Day-Carroll,L., Dedrich,D.A.,
        Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H.,
        Douthett,K.J., Drepper,H., Dugan-Rocha,S., Durkin,K.J.,
        Eathart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
        Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
        Gabist,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gall,I.R.,

```

```

CORRELL,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
HARRIS,C., HARRIS,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J.,
HERNANDEZ,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
HOMI,P., Howard,S., Huber,J., Huijck,S., Hume,J., Jackson,L.B.,
JACOBSON,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
KARLSON,E., Kelly,S., Khan,U., King,L., Kovach,C.,
KRATOVIC,J., Kureishi,A., Landry,N., Leal,E., Lewis,L.C., Lewis,L.,
LI,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Lounsbury,H.,
LOZADO,R.J., Lu,X., Lueker,A., Lueker,R., Luna,R., Ma,J.,
MAHSHARI,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
MASSEY,E., McWhinney,B., McLeod,M.P., Meador,M., Mei,G., Metker,M.,
MINER,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,N., Morris,S.,
MOSEY,N., Neal,D., Newton,E., Nwokenwo,S., Ogun,M., Okunolu,G.,
ORAGUNYE,N., Oyedero,R., Pace,A., Payton,B., Peery,D., Perez,L.,
PETERS,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
RIVERS,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
SODERGREN,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
SULTON,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
TANSEY,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
USMANI,K., Vasquez,L., Vera,Y., Villalon,D., Vinson,R., Wang,Q.,
WANG,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
WILLIAMS,C., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
WU,C., Wu,X., Wu,Y.F., Zhou,Y., Zorilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
2 (bases 1 to 161955)
Worley,K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 161955)
Worley,K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:18138893.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web Site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center Project name: GKTO
Center clone name: CH230-117N16
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 102133 bases at least Q40
Consensus quality: 108299 bases at least Q30
Consensus quality: 114299 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1108: contig of 1108 bp in length
* 1109 1208: gap of unknown length
* 1209 2485: contig of 1277 bp in length
* 2486 2585: gap of unknown length
* 2586 3731: contig of 1146 bp in length
* 3732 3831: gap of unknown length

```



```

Db      160554  GACAGGCAATTCATGACATGGCTGCAGATCAAAAAACGGATGGGCCAGAGCAGCA 160495
QY      150  AAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 169
      ::|||||  ||:::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      160494  TCTGGGGGT---GACGGGCCCAACCTGAGATCGACAGCACTCTCTGTGAAATCTGCTAGT 160438
QY      170  GLyGlyCysCys 173
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      160437  GGTGGCTGCTGC 160426

```

Search completed: June 22, 2003, 18:32:05  
 Job time : 1643 secs

